DOCUMENT RESUME

ED 176 989 SE 028 989

AUTHOR	Hord, Shirley M., Ed.; Hall, Gene E., Ed.
TITLE	Teacher Education Program Evaluation and Follow-Up Studies: A Collection of Current Efforts.
INSTITUTION	Texas Univ., Austin. Research and Development Center for Teacher Education.
SPONS AGENCY	National Inst. of Education (DHEW), Washington, D.C.
PUB DATE	78
NOTE	198p.; Contains occasional marginal legibility particularly in Tables and Figures
AVAILABLE FROM	University of Texas, Research and Development Center for Teacher Education, Austin, Texas 78712 (Publication 7000; No price quoted)
EDRS PRICE	MF01/PC08 Plus Postage.
DESCRIPTORS	Educational Programs; *Educational Research; Evaluation; *Followup Studies; *Inservice Teacher Education; *Preservice Education; *Program Evaluation; Science Education; Teacher Education

ABSTRACT

Presented is a collection of papers which describe follow-up studies designed to evaluate the effectiveness of some teacher education programs at six colleges and universities. Institutions represented are Western Kentucky University, Weber State College, University of Oregon, The Ohio State University, Tennessee Technological University, and the University of Houston. Ten papers are included in this publication since more than one presentation was made by representatives from several institutions. The final chapter of the document contains a summary of the studies as well as the identification of eight implications for future follow-up studies, research teaching, and research in teacher education. (PB)

****

* Reproductions supplied by EDRS are the best that can be made *

* from the original document. *

****

QB

RESEARCH AND DEVELOPMENT AGENDA IN TEACHER EDUCATION

ED176989

U S DEPARTMENT OF HEALTH
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Nancy Wilder

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TEACHER EDUCATION PROGRAM EVALUATION AND FOLLOW-UP STUDIES:

A Collection of Current Efforts

Editors:
Shirley M. Hord and Gene E. Hall

SE 028 989

Teacher Education Program Evaluation and Follow-Up Studies: A Collection of Current Efforts

Shirley M. Hord
Gene E. Hall

Editors

Research and Development Agenda in Teacher Education

The Research and Development Center For Teacher Education
The University of Texas at Austin

UTR&D

These papers originally shared at an initial colloquium (on teacher education program follow-up studies), hosted by the University of Texas R&D Center for Teacher Education, April 26-28, 1978, in Austin, Texas.

Additional information may be obtained from the authors whose names and addresses may be found at the end of this volume, "List of Contributing Authors."

The research described herein was conducted under contract with the National Institute of Education. The opinions expressed are those of the authors and do not necessarily reflect the position or policy of the National Institute of Education, and no endorsement by the National Institute of Education should be inferred.

Publication #7000 available from
the Research and Development Center for Teacher Education.

TABLE OF CONTENTS

Page

FOREWARD
Gene E. Hall, Research and Development Center for Teacher Education, The University of Texas 1

I. WESTERN KENTUCKY UNIVERSITY FOLLOW-UP EVALUATION OF TEACHER EDUCATION GRADUATES
Ronald D. Adams, Western Kentucky University 11

II. EVALUATION OF THE WEBER STATE COLLEGE CBTE PROGRAM
Harley K. Adamson, Caseel D. Burke, David R. Cox, Weber State College 39

III. SECONDARY TEACHER PREPARATION PROGRAM UNIVERSITY OF OREGON PROCEDURES FOR PROGRAM EVALUATION
Richard I. Arends, University of Oregon 53

IV. THE OHIO STATE UNIVERSITY FOLLOW-UP STUDY
Judith D. Aubrecht, Kevin Ryan, Mikell O'Donnell, The Ohio State University 67

FOLLOW-UP STUDIES IN SCIENCE EDUCATION AT THE OHIO STATE UNIVERSITY
Patricia E. Blosser, Robert W. Howe, The Ohio State University 85

V. TEACHER EDUCATION PROGRAM STUDY AT TENNESSEE TECHNOLOGICAL UNIVERSITY
Jerry B. Ayers, Tennessee Technological University 99

VI. EVALUATIVE STUDIES OF GRADUATES FROM TEACHER PREPARATION PROGRAMS AT OREGON COLLEGE OF EDUCATION
H. D. Schalock, The Teaching Research Division, Oregon State System of Higher Education,
J. H. Garrison, G. R. Girod, K. H. Myers, Oregon College of Education 111

VII. EVALUATION OF INSTRUCTIONAL SYSTEM CHARACTERISTICS OF THE PROFESSIONAL TEACHER PREPARATION PROGRAM AT THE UNIVERSITY OF HOUSTON
Wilford A. Weber, James M. Cooper, University of Houston . 133

Page

PROFESSIONAL TEACHER PREPARATION PROGRAM EFFECTIVENESS STUDIES: 1976-1977 AFFECTIVE TESTING
Howard Jones, Robert Randall, University of Houston . . . 155

PTPP SUCCESS STUDY OF STUDENT TEACHERS
Howard Jones, Robert Randall, University of Houston . . . 169

SUMMARIZING THE STUDIES
Shirley M. Hord, Research and Development Center for Teacher Education, The University of Texas 185

LIST OF CONTRIBUTING AUTHORS 197

FOREWORD

In the 1970 standards adopted by the National Council for Accreditation of Teacher Education, it is stated that "systematic efforts to evaluate the quality of its graduates" should be a characteristic of teacher education programs as a basis for certification. In the 1979 NCATE standards emphasis on this concept continued with reference to inclusion of "evidence of their (graduates') performance relating to program objectives."

As is clearly indicated in the NCATE standards and other recent writings the profession is directly concerned about the performance of graduates of teacher education programs and is interested in reviewing data that assess the efficiency and effectiveness of teacher education programs. The profession is also interested in seeing data from follow-up studies of the graduates of these programs. Although this concern is clearly documented, the actual collection of data is spotty. The efforts to date to evaluate teacher education programs are very limited and surprisingly small in scale.

During the spring of 1978, several of us at the Texas R&D Center and some of our colleagues at the National Institute of Education shared a common interest in attempting to pull together what is presently known and understood about the design and conduct of teacher education program evaluations and follow-up studies. We contacted our teacher education colleagues at collaborating institutions as well as others around the country who should have some idea about the present state of the scene. We began with a list of those we knew to be seriously involved in teacher education program evaluations and follow-up studies and expected to be able to add to this list. Much to our disappointment we were unable to unearth a large number of institutions where there has been a serious commitment to conducting these studies. In order to shed more light on what presently is going on in a few institutions and in order to facilitate an opportunity to share experiences, the R&D Center and NIE hosted a three-day invitational meeting during April of 1978. The "doers and decision-makers" who have been involved in conducting teacher education program evaluation and follow-up studies were invited. This monograph is an outcome of that meeting.

The papers were authored as the ticket of admission to the meeting and represent the thinking of the authors as they were completing another year of data collection. They were asked to respond to a standard set of questions which were designed to provide information for discussion purposes at the conference. We also expected that the collection of papers could serve as a resource to other individuals and institutions as they become interested in designing and conducting program evaluation and follow-up studies.

The key points and questions that the paper authors were asked to address were:

1. Provide an overview description of your study (goals, objectives, research questions, design).
2. What is the history of the study (length of existence)?
3. Describe the measures (attach samples).
4. Describe your data base (what is in it? formats, examples).
5. What are you doing for feedback (to whom, how, with what effects)?
6. What are three or four key findings and their implications?
7. What are some problems you are presently encountering?
8. What are your research plans for next year?
9. Please cite the key references to your study.

The papers and the studies that were identified are interesting from several points of view. The studies are being conducted at a range of institutions from large state universities to small regional colleges. They represent a great deal of institutional and individual commitment to program evaluation and follow-up. In several cases, such as Western Kentucky University and Tennessee Technological University, the studies represent five or more years of longitudinal data that have been collected on graduates of their programs. In two institutions, researchers had just completed the first year of data collection on a pilot basis. They were being confronted first-hand with the realities of conducting longitudinal studies, including the disappearance of the samples and the challenges of attempting to justify traveling a long distance to collect data from an N of one! Other methodological difficulties are reported such as attempting to develop statistical inferences and generalizations based on small and decreasing samples.

However, in all cases, the authors and their papers represent extensive "real-world experiences."

Another interesting characteristic of these studies is that in all cases they have been conducted with individual and institutional resources and not with federal dollars. All of the institutions have been able to assign resources, rather modest in terms of their dollar value, that have resulted in extensive high quality data and key findings. In general, these findings have been of most use to the program developers and researchers at their own institutions and have not been shared as widely as perhaps they should have been. In many cases the studies have provided an opportunity to develop research expertise and an avenue for publication for those who have invested in the studies.

The papers are offered here to help stimulate discussion around a activity that the profession is concerned about and expects all institutions to do as a basis for certification. At the same time, they represent documentary evidence that it is possible to conduct these kinds of studies with modest institutional resources and that the data can be of interest to evaluators, program developers, and researchers alike. The authors are most willing to serve as a resource and provide advice as well as to share war stories about the experiences they have had.

In brief overview, the chapters are outlined as follows:

In the first chapter, Ron Adams describes the TPEP evaluation program at Western Kentucky University. This evaluation program has a 7-year history of data collection and represents one of the first and most comprehensive of the program evaluation efforts to be presented. The studies have systematically followed a sample of graduates from each year of the program. Not only has demographic data been collected but direct observation of teachers' performance has been done during succeeding years of inservice. Among the findings are descriptions of the comparisons of variables for teachers who stay in teaching versus those who do not begin or do not stay in teaching. Adams also includes a description of teaching behavior across years of inservice and the relationship of inservice teacher behavior to various characteristics of the teacher education program. The TPEP evaluation model is based upon an AACTE monograph, authored by J. T. Sandefur, which outlined a possible model for the conduct of evaluation of teacher education programs.

The second chapter was prepared by Harley Adamson, Caseel Burke, and David Cox and describes the planning and evaluation activities of the teacher education program at Weber State College. Weber State College was one of the first institutions to make an institutional commitment to performance-based teacher education. They began operation of their first PBTE program in the fall of 1970. The first comprehensive evaluation of this program was initiated in 1973. The evaluation plan was based on a four phase design of which only the first phase was fully implemented. The basic design proposed a series of evaluation activities followed by program revisions and a new round of evaluations. The first phase of the evaluation design is heavily focused on the delivery system; student performance and opinion data were collected in relation to the instructional modules (WILKITS). The paper concludes with a brief description of the planning to implement Phase IV of the evaluation design, which would be a comparison across programs.

Evaluation of the secondary teacher preparation program at the University of Oregon is described by Richard Arends in Chapter Three. This data collection effort began during the spring of 1975 and was expanded in 1976. The purpose of the study is to make judgements about the effectiveness of the program and to use the findings from the evaluation activities to guide program development and revision. The study's audience ranges from prospective education majors who are considering entering the secondary education program, to a regional cosortium which has responsibility for program plans and approval for certification of University of Oregon programs. The Secondary Education Plan looks at graduate satisfaction, competency and what happens to graduates one year after graduation. Data collection included classroom observations and interviews with principals and teachers.

Chapter Four represents the first of two chapters in which there are major sub-sections representing different evaluation activities that are going on in large university settings. The Ohio State University is large and diverse in terms of program emphasis, staff interest, and capabilities. Therefore, several relatively independent yet complimentary teacher program evaluation efforts have been underway. Thus, in this chapter two different papers are reported, one by Judith D. Aubrecht, Keven Ryan, and Mikell O'Donnell and another by Pat Blosser and Robert Howe.

The Aubrecht, Ryan, and O'Donnell paper begins by pointing out that there are twenty-five different undergraduate teacher certification programs at Ohio State University! The evaluation plan that is described is an attempt to collect data across these different certification programs by studying a few of them each year. The first round of data collection was during the 1977-78 school year with 1977 graduates being the first to be followed. Each graduate was mailed a survey and then a small sample of graduates was observed. The basic plan for the study is based on early involvement of faculty in terms of their having input into the design stages of the evaluation plan and also in their identifying a consensus set of characteristics, behaviors, and attitudes that they believe their graduates should have. This paper also addresses many of the research design issues that are faced in the first year of an evaluation effort.

The second paper in Chapter Four, by Blosser and Howe, describes the continuing efforts at conducting evaluations and follow-up studies of graduates of the science education programs at Ohio State University. The paper describes efforts that took place between 1964 and 1976 when faculty and graduate students conducted a series of studies of undergraduates during preservice and follow-up studies of graduates. The studies have covered a range of evaluation questions and activities but all have been designed to examine strengths and weaknesses of the science education program, to determine the style of graduates and to explore possible effects of the school setting on teachers' success. Directions that program revisions should take are then considered. The studies cover a wide range of variables, although a common set of measures have been used across most studies. Key findings and implications, as well as some of the problems that have been encountered, are described.

In Chapter Five, Jerry Ayers describes the evaluation program at Tennessee Technological University. This program is another of the rare occurrences where there has been serious ongoing institutional commitment to evaluation of teacher education programs and conduct of follow-up studies. Through 1969 the university conducted mail surveys of its graduates.

Beginning in 1970 special studies and a more systematic survey of graduates was initiated. Then in 1973 the Tennessee Technological University Teacher Evaluation Model was begun. In this model all graduates from the past twelve months are surveyed during the fall. A longitudinal study of graduates teaching within a 100-mile radius of the university is conducted. This study also includes evaluation staff spending a day with each teacher. A combination of teacher interviews, teacher surveys, classroom observations, and student data are collected.

In Chapter Six, Del Schalock, J. H. Garrison, G. R. Girod, and K. H. Myers describe the evaluation studies of graduates at the Oregon College of Education. During the early 1960's there were scattered attempts to conduct follow-up surveys of graduates; however, this activity was dropped late in that decade. In 1972 the experimental beginning of a competency-based and field-centered approach to the preparation of elementary teachers was initiated at the Oregon College of Education. With the coming of this program there was a need for more systematic data which led to an informal evaluation study during 1972-73. Beginning in 1974-75 more systematic studies were conducted. An interesting feature of these studies is that constrasting methodologies have been tested. This paper includes samples of findings that result from use of these different methodologies -- mail only, telephone-mail survey, and on-site visitation. The paper also includes research cost estimates. During the years since 1975 the evaluation plan has become more systematic and has included data from secondary as well as elementary preparation programs. This paper concludes with a brief description of how the work at this institution is being shared and coordinated with what is happening elsewhere in the state of Oregon.

Chapter Seven is composed of three papers that describe the program evaluation and follow-up study activities that are being developed at the University of Houston. In the first paper, Will Weber and Jim Cooper describe the efforts of a task force to evaluate the instructional system which is being developed at the university. This instructional system is an outgrowth of the commitment that was made to develop innovative approaches to the delivery of professional teacher education training that dates back to 1966 and was culminated in the spring of 1977 with the faculty vote to incorporate competency-based approaches in all undergraduate instruction.

In the second paper, Howard Jones and Robert Randall describe the evaluation efforts that were conducted during the spring of 1977 to look at affective dimensions of the undergraduates in teacher preparation at the University of Houston. A series of tests that focused on assessment of self-image, attitude towards teaching, motivation, and cognitive style were administered by faculty in regular courses on a volunteer basis. The authors worked with faculty in interpretation of data and feedback were offered as quickly as possible to all students who had participated. The data were then aggregated by the various program emphases and compared. Similarities and differences across programs and descriptions of undergraduates in general at the University of Houston are reported.

Jones and Randall, in their second paper, report on a survey of the perceived success of student teachers during student teaching. A survey form was developed which assessed 15 of the 16 generic teaching competencies that are a part of the University of Houston undergraduate program. Then this survey was administered to student teachers, their university supervisors and the school-based teacher educator with whom the student teacher was assigned. Ratings of perceived importance of the 15 generic teaching competencies as well as the rating of success of student teachers within and across groups is reported, compared, and contrasted.

In the final chapter, Shirley Hord of the Texas R&D Center for Teacher Education, does an analysis and critique of the studies. She does this by identifying common themes, issues and findings from across the studies. She concludes her synthesis with the identification of a set of implications and issues. She also identifies targets for future research and makes suggestions for the design of future program evaluation and follow-up studies.

In these papers the authors have shared their problems and activities as well as their findings. The authors, and we at the Texas R&D Center, invite the reader to learn from these papers and to contact any of us for further information or an update on activities. As of this writing most of the institutions that are reported on are continuing to support the studies.

However, as several authors noted there are key policy decisions pending in some institutions which may affect the direction of continuation as well as perhaps the viability of these studies. In general, it seems that the future promises expansion of activity in the area of teacher education program evaluation and follow-up studies. The profession is likely to increase pressure for the data, and if the profession does not succeed in encouraging and facilitating this development, the legislative-public is certain to force it. We hope that these papers will provide others who become interested in conducting teacher education program evaluation and follow-up studies with useful information and help to prevent their having to rediscover all of the same problems that have been dealt with by the authors of these papers.

Gene E. Hall, Program Director
The Research and Development
Center for Teacher Education
The University of Texas at Austin

January, 1979

WESTERN KENTUCKY UNIVERSITY
FOLLOW-UP EVALUATION
OF TEACHER EDUCATION GRADUATES

Ronald D. Adams
Western Kentucky University

WESTERN KENTUCKY UNIVERSITY FOLLOW-UP EVALUATION OF TEACHER EDUCATION GRADUATES

Ronald D. Adams
Western Kentucky University

Overview

The Teacher Preparation Evaluation Program (TPEP) at Western Kentucky University is an ongoing, systematic follow-up of graduates designed to provide objective and quantifiable information from a variety of sources. Data are obtained through classroom visits to teachers who graduate and enter the teaching field.

This evaluation system provides for a sample of students to be selected each year and followed in subsequent years as they continue to teach. Participants are first observed as undergraduate student teachers, and again at the end of their first, third, and fifth years of teaching. Each year begins a new cycle of the evaluation, and each cycle consists of four phases. Phase 1 concentrates on evaluating student or preservice teachers, and subsequent phases evaluate the same participants as inservice teachers. Figure 1 illustrates this procedure.

CYCLE I	Phase 1	Phase 2		Phase 3		Phase 4	
CYCLE II		Phase 1	Phase 2		Phase 3		Phase 4
CYCLE III			Phase 1	Phase 2		Phase 3	
CYCLE IV				Phase 1	Phase 2		Phase 3
CYCLE V					Phase 1	Phase 2	
CYCLE VI						Phase 1	Phase 2
CYCLE VII							Phase 1
	1972	1973	1974	1975	1976	1977	1978

Figure 1 - Cycle and Phase Arrangement for TPEP

Cycles I through V utilized a random sample of forty participants taken from the spring semester of student teachers. Stratification was based on the type of certification sought, elementary or secondary, with twenty subjects randomly selected from each area. The student teacher participants comprised Phase 1 of each cycle. Participants for Phases 2, 3, and 4 of each cycle are Phase 1 participants who were employed as teachers and remained in teaching for five years, respectively. Information is obtained each year regarding participants' teaching status.

Cycles VI and VII have employed slightly different approaches to the selection of student teacher participants in an effort to increase the initial sample size and the first year teacher sample. These two samples were larger than forty, but were not randomly selected. Also, supervisors of student teachers were trained to collect much of the data for these samples. Data on previous samples were primarily collected by trained graduate assistants.

Data obtained from TPEP constitute a comprehensive data base from which various data analyses are performed. Analyses range from simple descriptive summaries to more complex multivariate analyses. Individual results are available only to the participant at the end of the fifth year.

History of the Program

Western Kentucky University's evaluation program is an outgrowth of a monograph written by J. T. Sandefur and published by the American Association of Colleges for Teacher Education (AACTE) in 1972. In the monograph, entitled An Illustrated Model for the Evaluation of Teacher Education Graduates, Sandefur advocated a systematic follow-up of teachers to obtain data on selected variables determined from the research literature on teacher effectiveness.

Planning for the implementation of the model was begun in the fall of 1971, with data first being collected during the spring of 1972. At the end of spring semester 1978, TPEP will have completed seven years of data collection. Data have been collected from 263 student teachers, 105 first year teachers, 40 third year teachers, and 9 fifth year teachers. Data are presently being collected from 61 student teachers, 40 first year teachers, 15 third year teachers, and 9 fifth year teachers.

There have been some modifications in TPEP over the years. Changes through 1976 were limited to instrumentation revisions and additions. In 1977 and 1978, major changes were implemented in Phase 1. Supervisors of student teachers were trained to collect data from a portion of student teachers. This year, in 1978, supervisors are again collecting data from 30 student teachers, with partial data being collected from an additional group of 31 student teachers. Data collection procedures for first, third, and fifth year teachers have not changed. Table 1 contains the specific changes made over the seven-year period.

An important consideration in longitudinal follow-up is the changes that occur which reflect current advances in the state of the art in educational research, yet maintain a data base that does not lose its longitudinal value. TPEP has attempted to make these changes in an efficient and practical way. Presently there are 200 plus variables, with repeated measures available for study. A more complete history of this program is available in the technical report entitled "Western Kentucky University's Teacher Evaluation Program: A Brief History from September 1971 through November 1977."

Measurement

Instruments utilized to measure the TPEP variables can be classified into four categories: career base line data; direct classroom observation; pupil, peer, and supervisor evaluations; and standardized measures. Instruments have been added at various stages of TPEP development. Table 1 illustrates this developmental process. Each category of data and the instruments used to obtain the data are briefly described:

Career Base Line Data

Career base line data for participants are obtained primarily from three sources--the Career Base Line Data Questionnaire, the Teacher Preparation Evaluation Inventory (TPEI), and transcripts of grades. A brief discussion of each instrument follows.

Career Base Line Data Questionnaire (CBLD). The CBLD questionnaire is

Table 1

Evolution of TPEP Procedures

INSTRUMENTS	72	73	74	75	76	$77^{@}$	$78^{\&}$
Career Baseline Data	I /	C / I	R / R	R / R	R / R	C / C	C / C
Grade Point Averages	I	C	C	C	C	C	C
F-Scale (45-40)	I	C	C	C	C	C	C
F-Scale (Balanced)		I	C				C
Rokeach Dogmatism Scale				I	C	C	C
TEP/S 1	I	C	C	R	R	C	C
TEP/S 2					I	C	C
Classroom Observation Record	I	C	C	C	C^{*}	C	C
Interaction Analysis	I	$C^{\#}$	$R^{\#}$	$R^{\#}$	$C^{\#}$	C	C
Student Evaluation of Teaching I	I	C	C	C^{+}	C	C	C
Student Evaluation of Teaching II			I	C^{+}	C	C	C
Teacher Preparation Eval. Inventory	I /	C / I	C / C	C / C	R / R	C / C	C / C
Observer Checklist					I	C	C
Teacher Concerns Checklist					I	C	C
Follow-up Questionnaire	I	C	C	C	C	C	C

Notes:
@ In 1977, supervisors of student teachers collected about ½ the data for Phase 1 Cycle VI.
* The factor analysis of the COR yielded 4 factors presently being used.
# IA ratios were added to data set.
+ The SET 1 & 2 forms were changed to allow computer scoring. Computer card forms and mark reader forms were utilized to aid scoring.
& In 1978, supervisors of student teacher collected about ½ the data for Phase 1 Cycle VII with partial data collected for the remainder.

Ledger: I = Implemented R = Revised C = Continued designed to obtain demographic data from participants that are not readily available from other sources. Initial personal and professional background information is collected for student teachers and updated as participants continue in the program.

Teacher Preparation Evaluation Inventory (TPEI). The perceptions of participants toward the university's preparation program are obtained through administering the Teacher Preparation Evaluation Inventory. In addition, measures of problems encountered in teaching are obtained by this instrument. Participants are asked to respond to a Likert-type scale for a number of items, as well as to four open-ended items.

Transcripts of Grades. Four grade point averages (GPA) are computed from the participant's transcript of grades. Overall GPA, professional education GPA, subject GPA, and student teaching grade are entered into the data files as separate variables at the completion of the student's undergraduate teacher preparation program.

Direct Classroom Observation

Two direct classroom observation systems are employed to obtain data on participants' classroom behavior. Observers are trained in the use of these systems, and interobserver reliability is established each year.

Classroom Observation Record (COR). The Classroom Observation Record, developed by Ryans, is used to assess four dimensions of pupil behavior and eighteen dimensions of teacher behavior. Each dimension is carefully described and defined in a glossary accompanying the recording form. A seven-interval scale is used to rate each of the pupil and teacher dimensions immediately after each observation period.

Sixteen of the eighteen teacher behaviors were found to form three major factors.* Factor A can be described as organized, confident, and systematic classroom behavior; Factor B as empathetic, understanding, and adaptable classroom behavior; and Factor D as stimulating, original, and alert teaching behavior. A fourth factor, Factor C, is composed of ratings of student behaviors characterized by alert, responsible, and confident classroom behavior. Two additional teacher dimensions not included in the factor structure are "fairness" and "attractiveness."

---

*The COR factor structure reported here closely approximates that found by Ryans.

Interaction Analysis (IA). A nineteen category interaction analysis system provides data on teacher-student verbal interaction and to a lesser extent, nonverbal behavior. This system is a combination of Flanders' and Hough's systems of interaction analysis. Twelve categories of teacher talk, three categories of student talk, and four nonverbal categories comprise the system. The observer records a numerical value corresponding to a specific category every three seconds or every time the category changes. Thus, an objective record is obtained of verbal interaction occurring in the classroom. Two twenty minute observations per participant are recorded.

Pupil, Peer, and Supervisor Evaluations

Perceptions about each participant are obtained from pupils, peers, and supervisors. These data are collected for each phase of TPEP with the only exception occurring in Phase 1. For Phase 1, the cooperating teacher, rather than peers and supervisors, rates the participant as a student teacher.

Teacher Evaluation by Peer/Supervisor I and II (TEP/S). Peers and supervisors are asked to rate participants on two instruments, TEP/S and TEP/S II. The first instrument was developed at Kansas State Teachers' College and obtains ratings on four broad items related to teacher behavior. The second instrument was developed by Adams and consists of seventeen items designed to measure three areas of perceived teacher behavior. The TEP/S II has recently been developed, and limited data are available at this time. For additional information on data here and elsewhere, write the authors at the address given in the "List of Contributing Authors."

Student Evaluation of Teaching I and II (SET). Perceptions of pupils about each participant are obtained from the SET I and II. SET I, developed by Veldman and Peck, allows students to rate the teacher on ten items that measure five dimensions of teaching behavior. This instrument was derived from the Pupil Observation Survey Report developed by McClain. Veldman found that SET I could be used with pupils in the third grade and above. However, use of this instrument has been limited to grade four and above for the TPEP study.

The Student Evaluation of Teaching II, developed by Haak, Kleiber, and Peck, measures three dimensions of teaching as perceived by pupils. True-False reponses are obtained from 23 statements regarding the teacher's behavior and feelings toward students. Two forms of this instrument allow data to be collected from pupils in kindergarten through grade six. A verbal, non-pencil-paper, card form is administered to kindergarten through third grade students, while a machine-scorable form is administered to fourth through sixth grade students.

Standardized Measures

Three instruments are administered to participants to obtain measures of authoritarianism, dogmatism, and level of concern about teaching. These data are collected for each Phase of TPEP.

F-Scale. The F-Scale, forms 45 and 40, was developed by Adorno and others to measure individual prejudices and antidemocratic tendencies. This 28-item scale refers to opinions regarding a number of social groups and issues.

Rokeach Dogmatism Scale. To provide supplementary data regarding dogmatic tendencies, the Rokeach Dogmatism Scale is administered to each participant along with the F-Scale. This 40-item scale measures dogmatic tendencies by requiring the participant to respond to a six-point, agree-to-disagree scale for each item. The Rokeach scale was added in 1975, and data are incomplete for earlier participants.

Teacher Concerns Checklist. A recent addition to the instrumentation of TPEP is the Teacher Concerns Checklist. This instrument allows for measures of the concerns teachers have about the teaching profession and about students. The first administration of this instrument in TPEP was during the spring 1976 data collection period.

Collection of Data

Observations are made toward the end of the spring semester with two visits scheduled per teacher. Both observations are made of the same class and at the same time of day. A twenty-minute interaction analysis recording and ratings utilizing the Classroom Observation Record are obtained at each visit.

The appropriate Student Evaluation of Teaching form is administered at the second observation period. The last fifteen minutes of the class period are requested for this administration. The Teacher Evaluation by Peer/Supervisor is given to the appropriate personnel during the first visit and is collected at the second visit. Instruments to be completed by the participating teachers are presented at the first visit and collected at the second visit. After all instruments have been collected and scored, data are placed in the participants' individual files to await transfer to computer files. Care is taken to protect the confidentiality of all data collected.

Management and Analyses of Data

Once data have been scored by research personnel, optical scan equipment, and/or special computer programs, data files are built and placed on disk for analyses. The magnitude and complexity of the data set and the need to maintain individual integrity of each phase and cycle requires a flexible data management system. The OSIRIS data management system, developed by the University of Michigan, affords flexibility in data management and is compatible with the Statistical Package for the Social Sciences (SPSS), a widely used set of statistical programs.

Data Base

The TPEP data are stored, managed, and analyzed with the help of an IBM 370-model 165 computer and selected software packages. Data are stored via two OSIRIS type files: one contains only student teacher data and a second, master file, contains the follow-up data. SPSS, OSIRIS, BMD, and locally developed statistical packages provide a wide variety of analysis techniques to which TPEP data may be subjected for analyses.

A matrix arrangement of variable number by subject ID number allows for maximum flexibility in management and analyses. By utilizing these two "controls" and filtering techniques available through OSIRIS, subject groupings are relatively easy to arrange while maintaining the integrity of the master data file.

An important consideration in this program is the ability to machine score and computer score instruments, especially the Student Evaluation of Teaching I and II and the Interaction Analyses. This ability saves many hours and allows for a reasonable time frame in data processing.

Feedback

Various forms of feedback have been attempted, ranging from prepared reports to each faculty member in the College of Education, to invitations for special analyses, to utilizing faculty members to collect data. It has been the objective of TPEP to provide information for programmatic decision-making to faculty members as well as to administrators. The results of dissemination efforts have been somewhat disappointing to date.

The first annual report was presented to faculty of the College of Education and consisted of a description of procedures and measurement, together with a summary of data collected that year. An invitation was issued to the faculty to do additional analyses using the data set. Little response was obtained to either. Additional reports have been made available to departments over the years. Only one report has been somewhat successful. Each year TPEP participants have been asked to respond to four open-ended questions. These responses are recorded and sent to each department each year. Faculty members seem to be more favorable toward this type of feedback.

In 1977, several faculty members were asked to participate in data collection for Phase 1, student teachers, as part of their supervision responsibilities. Seven faculty members participated in this program. While successful in the data collection, little interest in the data analyses has been evident.

The feedback efforts to date have been of a technical nature, usually in report form with statistical tables and graphs followed by findings written in typical research jargon. Little opportunity has been provided for findings to be translated into programmatic implications. It may be that for meaningful feedback to occur, evaluators and program faculty members must work together in interpretation of data and in eventually deciding what programmatic changes should occur as a result of those interpretations.

Findings

The analyses of TPEP data are just beginning to provide interpretable, meaningful information; in large part this is due to the increasing n-size and to replication of measurement. With the large number of variables collected in each phase, analyses must be problem oriented; that is, questions must be formulated and analyses performed to address those questions, as opposed to a shotgun approach to data analyses. The following findings are presented as examples of this problem-oriented approach.

Problem One: What are the factors related to perceived problems of first year teachers?

First, the rank order of problems as perceived by first year teachers was determined. Both secondary and elementary teachers perceived the severity of selected problems in much the same way. Table 2 gives this information. "Teaching disrespectful students" and "Discipline" ranked, respectively, as numbers one and two, with "Motivation of students" ranked third. These findings were not surprising; however, they did emphasize the need for preparing first year teachers to deal with behavioral problems of students at all grade levels.

The next analyses dealt with trying to determine variables related to problems perceived by first year teachers. The "problems" variables, from the Teacher Preparation Evaluation Inventory, were correlated with observed teacher behavior variables and supervisor/peer ratings. These analyses allowed for patterns of significant correlations at the .05 level to be determined. For the sake of brevity, only the results from secondary teachers will be reported.

The perceived problems for which patterns of correlation coefficients were found were "Teaching disrespectful students," "Teaching students of different socio-economic levels," "Discipline," and "Relevance of materials to students." It was found that teachers who demonstrate desirable teacher behavior, as observed by using the COR, perceived these same problems as less severe.

TABLE 2

Summary of First Year Teachers' Responses to the Seriousness of Selected Problems

Item	Description	Secondary			Elementary		
		MEAN	S.D.	RANK	MEAN	S.D.	RANK
1	Student-Teacher Rapport	3.17	.76	7.5	3.46	.74	10
2	Teaching Phys. Handicapped	3.64	.58	10	3.26	.75	7.5
3	Teaching Disrespectful Students	2.46	.65	1	2.88	.83	1
4	Teaching Students with diff. S-E levels	3.08	.72	5.5	3.26	.66	7.5
5	Discipline	2.50	.88	2	2.93	.81	2
6	Relevance of Materials	2.75	.74	4	3.10	.67	6
7	Faculty-Teacher Relationships	3.83	.38	12	3.74	.50	11
8	Administrator-Teacher Relationships	3.71	.46	11	3.78	.42	12
9	Parent-Teacher Relationships	3.50	.59	9	3.43	.63	9
10	Motivation of Students	2.58	.78	3	2.95	.76	3.5
11	Self-Evaluation of Teaching	3.08	.72	5.5	2.95	.73	3.5
12	Knowledge of Outside Resources	3.17	.56	7.5	3.07	.63	5

Secondary N = 24
Elementary N = 42

Teachers who were more indirect perceived less problems with student behavior. Teachers who were rated higher by supervisors and peer teachers on the variable "Student rapport" and "Student expectations" perceived less problems than did teachers who were rated lower on these items. Table 3, Table 4, and Table 5 contain the results of these analyses. The general conclusions that may be drawn from these data are that teachers who perceive severe problems also demonstrate less effective classroom behavior, and they are viewed by supervisors and peers as being less effective in the areas of student rapport and student expectations.

Implications of these findings for teacher education are not clearly known at this time. However, it would appear that first year teachers do vary in their teaching behavior and characteristics, and that their problems in teaching are related to these behaviors and characteristics. Peer and supervisor relations also seem to be related to first year teachers' problems. If these findings could be translated into teacher education programs, problems experienced by first year teachers could, perhaps, be lessened.

Problem Two: Identify probable factors related to teachers' entry into teaching and to teachers' retention after three years of teaching.

Recent studies have indicated that attrition of teachers entering and remaining in the profession is extremely high (about 66 percent, Joyce, 1977). These estimates are holding true at Western Kentucky University. It can be seen in Table 6 and Table 7 that the retention rate from student teaching to first year teaching was 70 percent for elementary and only 43 percent for secondary. After three years, the retention rate is 53 percent for elementary and 27 percent for secondary. At the fifth year, only 28 percent elementary teachers were still teaching and only 18 percent of secondary teachers were still teaching. These estimates are somewhat low because teachers returning to the profession were not considered, i.e., if they were not teaching during the data collecting phase, they were dropped from the study.

Information as to why they did not remain in teaching was collected from non-teaching participants. These data are contained in Table 8 and Table 9. While "could not find a job" was the most frequent reason given for not entering teaching, there is some reason to believe that the locale in which the teacher wished to teach was an important consideration. Between the first and

TABLE 3

Correlation Matrix of COR Measures and Four
Problem Areas as Perceived by
First Year Secondary Teachers

Item	Factor A	Factor B	Factor C	Factor D	Fair	Attractive
3 (Disrespectful Students)	.42	--	.49	.44	.36	--
4 (Diff. So-E Levels)	.41	.47		.35	.50	.62
5 (Discipline)	.35	--	.60	--	.38	--
6 (Rel. of Materials)					.49	.45

N = 24

TABLE 4

Correlation Matrix of Selected IA Ratios to Four Problem Areas
as Perceived by First Year Secondary Teachers

Item	i/d	I/D	Stu. talk/Tech. talk
3 (Disrespectful Students)	.48	---	---
4 (Diff. S-E Levels)	---	.42	.43
5 (Discipline)	---	---	---
6 (Rel. of Materials)	---	---	-.43

TABLE 5

Correlation Matrix of Supervisor and Peer Ratings and
Four Problem Areas as Perceived by First Year Secondary Teachers

	Supervisor Ratings of Tech.		Peer Ratings of Tech.	
Item	Student Rapport	Student Expectations	Student Rapport	Student Expectations
3 (Disrespectful Students)	.40	---	.42	---
4 (Diff. S-E Levels)	---	.42	---	---
5 (Discipline)	.36	---	.46	.38
6 (Rel. of Materials)	---	---	---	---

TABLE 6

Number of Teachers Who Remained in the Study Sample Over a Five-Year Period

Cycle	Level	Phase 1 Student Teaching	Phase 2 1st Year Teaching	Phase 3 3rd Year Teaching	Phase 4 5th Year Teaching
Cycle I (1972)	Elementary	20	15	10	6
	Secondary	18	7	3	3
Cycle II (1973)	Elementary	20	13	7	5
	Secondary	20	8	5	4
Cycle III (1974)	Elementary	20	16	13	
	Secondary	17	9	8	
Cycle IV (1975)	Elementary	17	14	11	No data for Cycle III Cycle IV Cycle V and Cycle VI
	Secondary	20	6	4	
Cycle V (1976)	Elementary	20	10	No data for Cycle V and Cycle VI	
	Secondary	20	7		
Cycle VI (1977)	Elementary	39	25		
	Secondary	32	17		
Total	Elementary	136	93	41	11
	Secondary	127	54	20	7

TABLE 7

Percent of the Original Study Sample Who Enter and Remain in Teaching in Kentucky

Time	Percent Teaching 1 Year after student teaching $N = \frac{145}{263}$	Percent Teaching 3 Years after student teaching $N = \frac{61}{152}$	Percent Teaching 5 Years after student teaching $N = \frac{18}{78}$
Level	Data collected on samples from six cycles	Date collected on samples from four cycles	Data collected on samples from two cycles
Elementary	70	53	28
Secondary	43	27	18
Combined Levels	55	50	23

*A participant who does not obtain a teaching position or moves out of state is dropped from the study.

TABLE 8

Percent of Participants Not Entering Teaching
Following Graduation by Reasons Reported

Reason for Not Entering Teaching	Percent of Responses		
	Elementary (N = 43)	Secondary (N = 73)	Total (N = 116)
Could not find a job	36	49	45
Graduate school	7	11	9
Marriage and pregnancy	5	5	5
Moved out of state	11	4	7
Did not want to teach	7	8	8
Other employment	4	5	5
Other reasons	7	3	7
Did not report	23	8	15

TABLE 9

Number of Participants Dropping Out of Teaching
Between Years One and Three by Reasons Given

Reason	Number Dropping Out		
	Elementary	Secondary	Total
Could not find a teaching position	2	0	2
Salary too low	2	0	2
Graduate school	0	2	2
Pregnancy	5	3	8
Moved out of state	2	0	2
Other employment	1	1	2
Did not report	5	4	9
Total	17	10	27

third years, pregnancy appears to be an important reason for leaving. It is not known if and when these teachers return to teaching.

Perhaps a more revealing picture of factors influencing this problem was obtained when the non-teaching group was compared to the teaching group. For secondary participants the largest attrition occurred between student teaching and the first year of teaching. Table 10 presents significant differences found via t-tests and chi square analyses for these two groups. Data from Table 10 suggest that members of the teaching group perceived fewer problems in establishing rapport with pupils and in being able to use alternative teaching methods during their student teaching experiences than the non-teaching group. The teaching group appeared to consider the quality of instruction or the sufficiency of instruction that they received in their teacher preparation program as greater than did the non-teaching group. In addition, the teaching group indicated that, for them, there had been more time spent during their training in the areas of teaching students of different levels and in developing relationships with the faculty.

Little differences were found between the groups in teacher behavior or characteristics. One exception was that the non-teaching group had more student-initiated talk occurring in their classrooms than did the teaching group. And, while not shown in the table, "Area of subject preparation" was found to be a good predictor.

Tables 11 and 12 present comparisons of teaching and non-teaching groups for elementary participants who dropped out of teaching between years one and three. The dependent variables in Table 11 were measured during their student teaching and dependent variables for Table 12 were measured at the end of their first year of teaching. Specific discussion of the findings will not be presented due to space limitations. However, it may be generally stated that student teachers who made better grades in professional education courses were observed as more competent, and had a better attitude toward the preparation program, remained in teaching. First year teachers who had higher peer and supervisor ratings, more positive attitudes toward their preparation program, and were less directive, stayed in teaching longer.

Problem Three: Does teacher behavior change with experience?

Repeated measures analysis of variance was utilized to determine if teacher

TABLE 10

Differences Between Secondary Student Teachers Who Obtained Teaching Positions and Those Who Did Not with Respect to Variables from Student Teaching Data

Variable	Instrument	Difference	Level of Significance
Perceived problem of developing rapport with students	TPEI*	Teaching group saw problem less severe	<.01
Sufficiency of instruction in teaching students with different ability levels	TPEI	Teaching group related the sufficiency of instruction as greater	<.05
Sufficiency of instruction in developing relationships with faculty	TPEI	Teaching group rated the sufficiency of instruction as greater	<.05
Perceived problems in the use of alternative teaching techniques	TPEI	Teaching group saw problem less severe	<.05
Quality of instruction in the use of alternative teaching techniques	TPEI	Teaching group rated the quality of instruction higher	<.01
Quality of instruction in tests and measurement	TPEI	Attrition group rated the quality of instruction higher	<.05
Ratio of student initiated talk to total classroom talk	Interaction Analysis	Teaching group .07 mean Attrition group .12 mean	<.05
Size of community in which participant was reared	TPEI	Participants reared in suburban areas more likely to obtain a teaching position	<.01
Membership in a professional organization during student teaching	TPEI	Over 50% of teaching group were members. Less than 25% of attrition group were members	<.01

*Teacher Preparation Evaluation Inventory

TABLE 11

Difference Between Elementary Teachers Who Were Teaching in Year Three and Participants Who Dropped Out of Teaching Between Years One and Three with Respect to Variables from Student Teaching Data

Variable	Instrument	Difference	Level of Significance
Professional education grade point average		Teaching group 3.46 Attrition group 3.12	<.05
Teacher appears stimulating as opposed to dull	COR*	Teaching group judged more stimulating	<.05
Teacher appears original as as opposed to stereotyped	COR	Teaching group judged more original	<.05
Teacher appears alert as opposed to apathetic	COR	Teaching group judged more alert	<.05
Teachers rating of their subject matter preparation	TPEI**	Teaching group rated their preparation as more adequate	<.05
Quality of instruction in developing rapport with students	TPEI	Teaching group rated the quality of instruction higher	<.05
Sufficiency of instruction in developing rapport with students	TPEI	Teaching group rated the sufficiency of instruction as greater	<.05
Quality of instruction in self evaluation of teaching	TPEI	Teaching group rated the quality of instruction higher	<.05
Usefulness of class in audio-visual teaching	TPEI	Teaching group rated the AV class as more useful	<.05
Perceived availability of faculty in the preservice program	TPEI	Teaching group rated the faculty as more available	<.01

*Classroom Observation Record
**Teacher Preparation Evaluation Inventory

TABLE 12

Differences Between Elementary Teachers Who Were Teaching in Year Three and Participants Who Dropped Out of Teaching Between Years One and Three with Respect to Variables from Year-One Data

Variable	Instrument	Differences	Level of Significance
Teachers average rating by supervisors	TEP/S*	Teaching group had higher ratings	<.05
Subject matter competence rated by peers	TEP/S	Teaching group had higher ratings	<.05
Teaching competence rated by peers	TEP/S	Teaching group had higher ratings	<.05
Quality of instruction in relevance of materials to students	TPEI**	Teaching group had higher ratings	<.05
Quality of instruction in motivating students	TPEI	Teaching group rated instruction higher	<.05
Sufficiency of instruction in developing faculty relations	TPEI	Teaching group rated instruction as substantial Attrition group rated instruction as excessive	<.05
Sufficiency of instruction in developing relationships with administrators	TPEI	Teaching group rated instruction as substantial Attrition group rated instruction as excessive	<.05
Ratio of indirect teacher talk to direct teacher talk	Interaction Analysis	Teaching group - .78 mean Attrition group- .46 mean	<.05
Ratio of lecture to total classroom time	Interaction Analysis	Teaching group - .16 mean Attrition group - .09 mean	<.05
Ratio of teacher talk to total classroom time	Interaction Analysis	Teaching group - .53 mean Attrition group - .44	<.05

*Teacher Evaluation by Peer/Supervisor
**Teacher Preparation Evaluation Inventory

behavior changed as a result of teaching experience. Data were available from Cycle I, II, and III participants with n-sizes of 29 elementary and 16 for secondary. Graphs in Figure 2 were constructed to describe the differences between Phase 1, 2, and 3 for observations utilizing the COR.

Secondary and elementary participants showed similar patterns of change over three years of teaching experience. A most dramatic change appeared to occur between Phase 2 and Phase 3, that is, between the first and third years of teaching. Change in interaction analysis variables was not found to be significant across years of experience. An implication of these findings for teacher education programs would be the need for more support of new teachers during their first year of teaching. It may be that new teachers entering the teaching profession need an induction period whereby they receive support from the teacher preparation institution as well as from the local education agency.

Data from one sample of elementary teachers (n=6) was complete for all four phases. While repeated measures analyses of the COR data were not statistically significant, they did reveal an interesting pattern. It may be that teachers during their fifth year of teaching may demonstrate less desirable teaching behavior than during the third year. If this trend were supported in subsequent samples, the need for additional inservice teacher education between the third and fifth years would be suggested.

Problems

Problems encountered in the conduct of Western Kentucky University's TPEP fall into four broad categories: (1) Management and Conduct of Data Collection, (2) Data Processing and Analyses, (3) Feedback Procedures, and (4) Research. Each of these areas is briefly discussed in the following paragraphs.

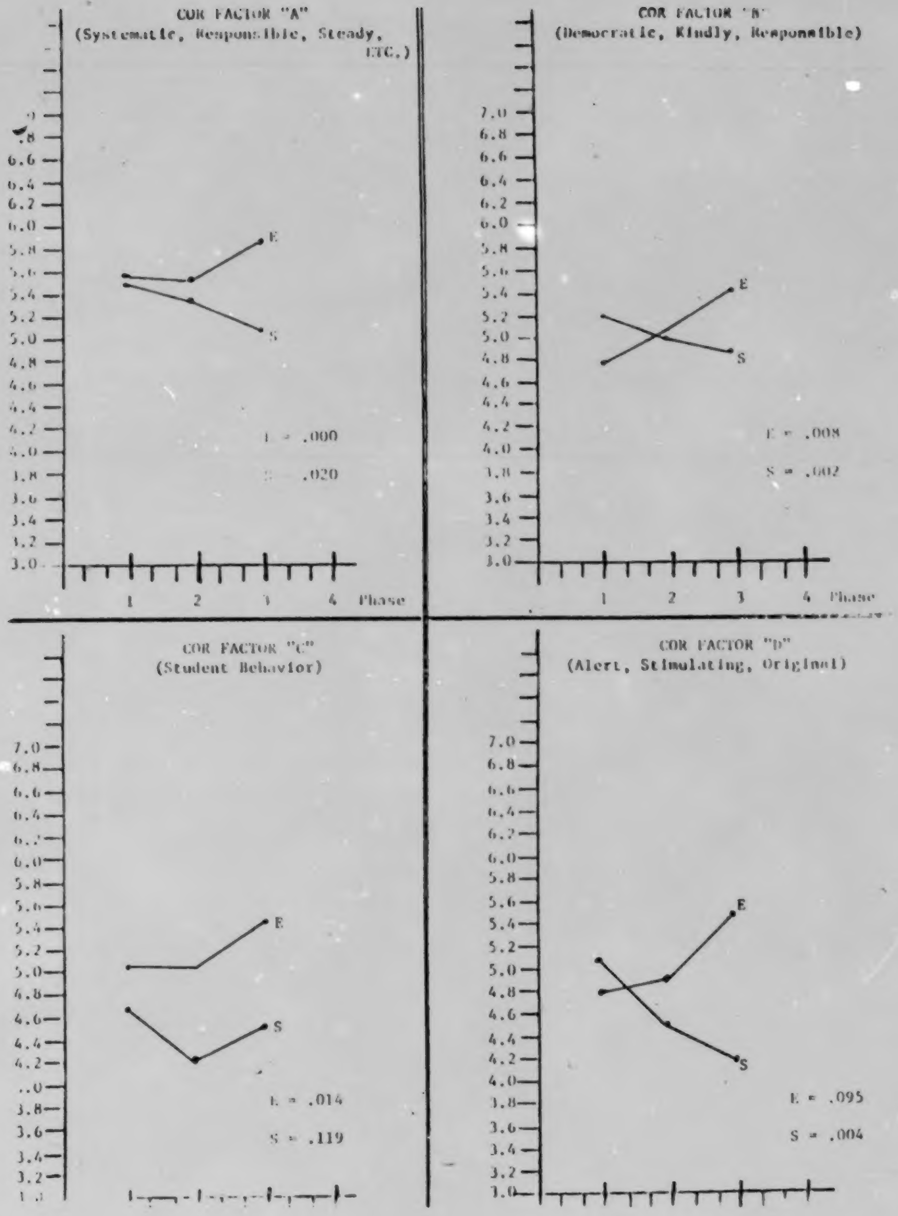


Figure 2 - Graphs of COR Repeated Measures ANOVA across TPEP Phases 1, 2, and 3.

Management and Conduct of Data Collection

1. Graduate assistants are the primary source of manpower for data collection. Since Western Kentucky University does not offer a doctorate, most graduate students finish their programs in one year, with the exception of psychology majors, who normally spend from three semesters to two full years in their graduate programs. This means that much time and effort must be spent in recruiting capable graduate assistants each year to work in TPEP. To compound the problem, observer training and data collection are conducted in the spring semester, and due to budgetary limitations, most of the graduate observers must be recruited during this time, after other assistantships have been filled.

2. Training of observers has been difficult due in large part to the difficulty in obtaining appropriate materials. Half-inch and three quarter-inch video tape units have provided most of the classroom situations used in training, supplemented with some films and live observation. The general lack of quality, quantity, and diversity of the materials have been a major problem in training observers.

3. The constant attention to record-keeping of the location of participants has become more of a problem as numbers of participants have increased. These data must be kept from year-to-year for each teacher. Presently this follow-up is done via mailed forms and telephone calls, a process that takes approximately two months.

4. The public relations aspect of the program is an increasing concern. Participants are sometimes hesitant to participate and often must be convinced of the program's importance. Also, the possibility of an incident occurring while an observer is in the school is always present. The protection to human subjects and right to privacy acts are sources of potential problem areas. With the increase in n-size projected over the next few years, these problems must be dealt with more efficiently.

Data Processing and Analyses

1. A major problem with data processing has been the acquisition and training of graduate assistants to handle the complex data sets required for TPEP. In most instances, graduate students do not possess the required computer skills and must be trained "on the job." The high turnover rate of graduate assistants also presents a transition problem in maintaining the data files. This problem has been diminished in past years by the services of a faculty research associate employed by project funds in another

administrative area. This individual has provided expertise in training graduate assistants and in developing custom software essential for efficient utilization of the computer. The future ability of TPEP to process, manage, and analyze data may rest with a more permanent position being established to work with data processing problems.

2. The opportunity to analyze data and disseminate findings has been limited. Resources allotted to TPEP afford just the time to do an adequate job of collecting and maintaining the data. Little remains for the in-depth analyses and report writing that should follow.

Feedback Procedures

The problem with feedback procedures have already been discussed. This is a critical issue in the survival of TPEP at Western Kentucky University.

Research

The validity of the evaluation model must be considered in long range planning. To what extent are the variables being measured reflective of teacher effectiveness? Recent research syntheses by Dunkin and Biddle (1974), Brophy and Evertson (1976), and Medley (1977) have established reasonably well that "effective" teaching is dependent upon the context in which it occurs. Such variables as age level and background (socio-economic status) of children and subject taught are factors in determining what type of educational experiences are most effective in student learning. Other equally important considerations are in pupil affective outcomes. Thus, what is appropriate teacher behavior for one teaching situation may not be appropriate for another situation.

The TPEP model utilizes a set of measures established to determine desirable teacher behavior/characteristics from review of the research in 1972. While these measures were shown to be factors in effective teaching, the appropriateness of variables should be studied in relation to pupil outcome measures within various school contexts. This type of research would allow intelligent decisions to be made on what variables can be eliminated and what new variables need to be added in evaluating teachers in differing contexts.

Such research is costly and beyond the resources available to Western Kentucky University of this time.

It should be noted that an initial effort to establish the feasibility of such a study was made in cooperation with a local Head Start district. Research was conducted utilizing pupil outcomes and teacher measures from seven Head Start classrooms. Findings from this study as they apply to TPEP were discussed in a recent report (Adams and Shiek, 1978).

Research Plans

Research plans for TPEP have not been finalized for 1979. Some changes are anticipated and will be dependent upon input from teacher education faculty and upon the available resources. In the director's opinion, the following areas are in need of review for possible change:

1. Collection of data earlier in the preservice program. Data should be collected on students as they enter the teacher education program. Demographic attitudinal data and/or psychological data may prove beneficial as predictive of program success. Other data on professional development should also be maintained as part of TPEP.

2. As already mentioned, research on teacher effects in various contexts utilizing TPEP teacher measures and pupil outcome measures is badly needed. Of course, this research is dependent upon external funding. Plans to acquire such funding are being made for 1979.

3. In-depth analyses of existing data are planned for next year with more emphasis on dissemination and communication of findings to faculty and to the educational community at large.

4. Few procedural changes are expected for Phases 2, 3, and 4. However, data collection from student teachers will probably be made by university supervisors and be limited in scope. However, a larger sample (n=40) is expected to be followed up the first year, thus increasing the n size in Phases 2, 3, and 4.

References

Adams, R. D. A model for the evaluation of teacher education graduates. Paper presented at American Association of Colleges for Teacher Education annual conference, February, 1978.

Adams, R. D. Follow-up and evaluation of teacher education courses - A pilot study. Final Report to NIE, August, 1974. (ERIC Document Reproduction Service No. ED 090173)

Adams, R. D. Western Kentucky University's teacher preparation evaluation model, Phase 1, Cycle 1: First annual report, 1972. (ERIC Document Reproduction Services No. ED 67562)

Adams, R. D. Western Kentucky University's teacher preparation evaluation program: A brief history from September 1971 through November 1977. Unpublished technical report, January, 1978.

Adams, R. D., & Lewis, T. Student teachers: Those who do and those who don't. Western Kentucky University, Issues in Education, 2(3), 1976.

Adams, R. D., & Sandefur, J. T. Factors influencing the perceived problems of first year teachers. Paper presented at the Mid-South Educational Research Association annual conference, November, 1976.

Adams, R. D., & Shiek, D. A. Relationships among pupil outcome measures and teacher characteristics/behavior for preschool disadvantaged children. Paper presented at American Educational Research Association annual meeting, April, 1978.

Pankratz, R., & Harryman, E. The Western Kentucky University evaluation model: Implications for career decisions. Paper presented at American Association of Colleges for Teacher Education annual conference, February, 1978.

Sandefur, J. T., & Adams, R. D. A case study of second-year teacher education graduates. Journal of Teacher Education, 1973, 24(3), 248.

Sandefur, J. T., & Adams, R. D. An evaluation of teaching: An interim report. Journal of Teacher Education, 1976, 27(1), 71-76.

Sandefur, J. T. An illustrated model for the evaluation of teacher education graduates. Washington, D. C.: American Association of Colleges for Teacher Education, 1970.

Shiek, D. A., & Adams, R. D. Factor analysis of the classroom observation record. Western Kentucky University, Unpublished technical report, April, 1978.

EVALUATION OF THE WEBER STATE COLLEGE CBTE PROGRAM

Harley K. Adamson, Caseel D. Burke, David R. Cox
Weber State College

EVALUATION OF THE WEBER STATE COLLEGE CBTE PROGRAM

Harley K. Adamson, Caseel D. Burke, and David R. Cox
Weber State College

In the period 1967-69, the Teacher Education faculty at Weber State College decided to change from a rather traditional approach to teacher education to a system referred to as the Individualized Performance-Based Teacher Education Program (IPT). In making the change, a number of basic decisions were made by the faculty, the most important of which were:

1. The curriculum shall be based on the concepts, skills, attitudes and appreciation considered essential to successful teaching by the faculty, the students, and the personnel in the schools served by the Teacher Education Program, and as determined from the literature.

2. The delivery system shall be changed from predominantly group lecture to a series of individualized instructional modules, each dealing with a carefully selected concept, skill or attitude.

3. Courses and instructional modules within the courses will be subject to modification and adjustment as indicated by constant input from participants in the system.

4. Change to a new system will be total and abrupt and will be the only plan for the preparation of teachers (elementary and secondary) at Weber State College.

5. The system will be both operational and experimental, in keeping with the belief that the new practices have promise, but that their worth must be proved through careful investigation.

6. Consultant help will be sought for support in specialized developmental and operational procedures outside the expertise of the resident faculty.

7. An effort will be made to obtain outside financial support considered necessary in accomplishing the development and evaluation of the system.

In keeping with these decisions, the new system was developed and has been in operation since the autumn of 1970. From that time until the end of 1978, records show that 668 elementary teachers and 738 secondary teachers have graduated and have been recommended for Utah teaching credentials.

The Weber State College program is characterized by the following main features:

1. The teaching competencies are defined as behavioral objectives to be met.
2. The assessment criteria are behaviorally stated with mastery levels defined.
3. Assessments take into account both student performance and knowledge.
4. Student rate of progress depends on demonstrated competency.
5. Traditional grading has been abandoned in favor of credit (CR) or not credit (NC).
6. Laboratory experiences, including student teaching, are carefully prescribed to provide a meaningful and ongoing relationship between theory and practice.
7. Students make the decisions on when, where, and how they study; on scheduling of interviews, seminars, or otherwise seeking faculty help; on when they are ready for performance testing; as well as the speed and rate of their progress.
8. Students also have a voice in what to study; the kinds and locations of field experiences; and the development and evaluation of curricula and procedures establishing performance criteria.

Initial Evaluation

The first comprehensive approach to evaluation of the Weber State Program began in January of 1973, and continued until August of the same year. The evaluation team was composed of selected faculty members from Teacher Education. This group was directed by an evaluation specialist not involved in Teacher Education, who had joined the faculty at Weber State College in 1972 as head of instructional development.

Because of the extensive involvement of faculty members, efforts were made throughout this study to insure objectivity in assessment design and procedures employed. Methods and instruments were identified, developed, and accepted by the team before they were used.

The evaluative design consisted of four phases planned to cover a period from two to five years. Phase I focused upon the basic delivery system and Phase II was concerned with clarification and revision of program assumptions, concentrations and sequences, as based upon findings from Phase I. Phase III was designed to evaluate the revised program growing out of the activities associated with Phase I and Phase II. Phase IV would involve comparative studies between the Weber State Program and other teacher preparation programs. The four phases are further outlined as follows:

Phase I: Evaluation of the basic delivery system

a. Selected group of trained faculty will examine each component of each WILKIT* in light of established criteria.

b. Adequacy of the WILKITS in operation will be determined by gathering data at the student user level.

c. Adequacy of the competencies will be provided in light of the in-field needs of our graduates (appropriate to reality demands).

Phase II: Clarification and revision of program assumptions

a. Examination of the assumptions and intrinsic qualities of the entire program.

b. Examination of consistency between all major aspects of the program.

c. Clarify or revise the program assumptions, concentrations and sequence among program components and program structure elements so that they are relevant to the in-field teacher.

Phase III: Evaluate the revised program based on activities in Phase I and II

a. Revise Phase I which will be implemented at the program level to determine the relevance of the program as a whole.

---

*Instructional modules used in the Teacher Education Program at Weber State.

b. Extend the scope of the evaluation activities of Phase I to determine the relevance of the program as an instructional system.

Phase IV: Conduct comparative studies between Weber State College Teacher Education and that of other institutions

Eight months were spent in Phase I of the evaluation. The other phases were never fully implemented. However, the data from Phase I were used to study the total program, and most of the modules were revised as a consequence.

Since the efforts in this evaluation were directed largely to Phase I of the plan, perhaps we should look at what was done and the findings. As previously indicated, the study was concerned with the basic delivery system. Instructional modules (WILKITS) were evaluated at three levels. First, the evaluation team examined components of each WILKIT to ascertain how well they met the operational, theoretical, and philosophical standards which program operators had set for themselves. The general purpose was to improve internal consistency. The following example identifies one component of a WILKIT, gives the possible evaluation objectives, and provides a pertinent item from the instrument.

Components	Possible Evaluation Objectives
Objectives	Determine whether objectives meet a specified set of criteria

(Excerpt from instrument)

1. Students judge objectives to convey demands clearly (90%--student agreement minimal criteria before revision).

At a second level, the adequacy of WILKIT components were evaluated by means of student ratings. The purpose of these ratings was to determine the extent to which these elements met the needs of students, in view of demands imposed in the evaluation of performance on each module. The following is one example from the second level of Phase I.

Components	Possible Evaluation Objectives
Experiences	Determine adequacy of experiences preparing student for performance measure

(Excerpt from instrument)

The learning experiences in this WILKIT are sufficient for achieving the objectives of the WILKIT.

Strongly Agree _________ Agree _________ Neutral _________
Disagree _______________ Strongly Disagree _______________

The third evaluation level of Phase I focused upon perceived adequacy of the competencies developed, as tested in the field by program graduates. This level dealt with the scope, priority, and relevance of competencies developed in the program. In addition to faculty, student, and graduate input, data were also gathered from cooperating teachers in the public schools from building principals, district administrative personnel, and from representatives of professional organizations.

The information gathered at this level pertained to how well program components actually met field conditions. That is, does the program prepare people to perform satisfactorily on the job? Also, are WILKIT specified competencies appropriate in scope, priority and relevance of required performance levels to the demands of the job? The following is an example from the third level.

Components	Possible Evaluation Objectives
Experiences	Instruction is relevant to preparation of performing actual field requirement (High degree of agreement required)

Evaluation Results from Phase I

Analysis of student performance and opinion data indicates that students generally reacted favorably toward the delivery system used in the program. Also, the graduates of the program felt that competencies learned seemed to be what they needed as classroom teachers. Each of these judgments by graduates was confirmed by cooperating teachers, school principals, and school supervisory personnel.

Students agreed that successful participation in this program tended to develop increased self-reliance, learning to allocate time, learning to plan work strategies, and general self direction.

Students felt that the most effective learning experiences were microteaching, field experiences, seminars with faculty members and specially prepared monographs. They thought the least effective learning experiences were repetitive reading and unsupervised peer teaching.

Students and graduates indicated overwhelming approval for behavioral objectives. The objectives were useful in focusing upon critical elements in the WILKIT, as a basis for self assessment, and guiding preparation for final assessment.

Problems found seemed to revolve around scheduling and logistics and the desire for more direct instruction from faculty members, as well as the desire for a greater variety of learning experiences.

Other Evaluation Efforts at Weber State College

Informal Assessments

Several assessments of an informal nature have been made from time to time, some of which were summarized by Burke in 1974. These investigations show that graduates of the WSC teacher education program have a high rate of employment in a time of teacher over-supply. Knowledge of this fact has been instrumental in influencing some students to enroll in the WSC program in preference to other programs.

There seems to be general agreement among students and faculty that the modular system requires greater effort than did the previous traditional-type program.

Students generally graduate from the system with an attitude of confidence, a commitment to teaching, and a readiness to accept employment. Also, they are variously described by their supervisors as enthusiastic, empathetic to their students, adaptable, and having readily available teaching skills.

Deficiences, often of a specific nature, are also mentioned by some supervisors, and these have been considered by the teacher education faculty in curriculum decisions.

Applying a Teaching Skills Inventory

During the winter of 1974-75 a teaching skills inventory was used to determine the comprehensiveness of the WSC modular system. This inventory, developed at Utah State University by Stone, et. al (1975), identifies five major teaching skills clusters and fourteen skill areas, not including the subject and personal dimensions.

Each of the 301 individual objectives of the WILKIT system, including those of the Interaction (human relations) Laboratory, were analyzed with relation to the skills inventory. This process focused attention on those skill areas that are most prominently dealt with in the WILKIT system, as well as on those that from the stated objectives may need more attention.

As an evaluative device, the skills inventory seems to pose some useful questions about the WILKIT system that may well be investigated. One such question is whether or not the content of the curriculum is correctly reflected in the stated objectives. Since the inventory does not deal with inconsistencies that may exist between the objectives and what really happens within the system, it would be advantageous, as a part of the total system evaluation, to initiate a study to explore this matter.

Another useful study would be conducted into the nature of the objectives. Some were identified by the inventory as being global in nature, others as fitting more than one skill area. It is not to be implied that some carefully devised global objectives cannot be desirable within a system, but it is evident the clarity and precision of the objectives influence the way, and the extent to which, they are achieved.

The Teaching Assessment Form

Although it was not conducted as a formal part of the evaluation program, an effort has been made to correlate actual teaching practices with distinct elements of the preservice program through the Weber Teaching Assessment Form. This form serves as a guide for the prospective student teacher, as an aid to the student teacher and cooperating teacher, and as the final student teaching rating form in the graduate's placement file.

The Teaching Assessment Form was developed over a period of nearly two years by the combined elementary and secondary education faculties at Weber State College. It attempts to utilize the familiar teacher evaluation form as a device for the improvement of student teaching performance in terms of observable behaviors and asks the rater to estimate the regularity with which the behavior occurs. For a copy of this form, write to the authors at the address given in the "List of Contributing Authors."

Cooperating teachers are encouraged to utilize the "CARS" (Cooperative Assessment and Resulting Strategies for Improvement) approach to help student teachers perform at the highest level of effectiveness. This approach is drawn from the Supervisors Process Model developed at the University of California, Santa Barbara (Boyan, et. al, 1972-73), which suggests focus upon specific teacher behaviors as targets for change and a systematic process for effecting change. The CARS system pinpoints behaviors which are characteristic of each item on the Teaching Assessment Form by means of explicit questions, identifies observational systems which may be used in assessing particular aspects of performance, and lists WILKITS (modules) through which the student teacher can review procedures to be used for improving performance.

Cooperating teachers are urged to complete all, or part, of the Teaching Assessment Form at an eary stage in their student teacher's experience as a means of setting goals for improvement. This also makes for increased confidence in completing the final form for inclusion in the student teacher's placement file.

The Interaction Laboratory

The Interaction Laboratory for Teacher Development is a very basic introduction to the area of interpersonal relations. It has been used as the initial on-campus experience in education at Weber State College since 1970. It was developed as a structured activity to allow college faculty with limited training in the area to provide meaningful experiences for their students. The Lab has been used widely on campuses other than Weber's and has been the subject of a number of formal and informal research studies.

All but one of these studies used the Tennessee Self Concept Scale (TSCS). Results of the studies using the TSCS are rather mixed, but it can be generally concluded that students who have completed the Interaction Lab show some improvement in self concept as measured by the TSCS. The studies generally found a significant increase in overall positive response, but failed to find significance in the ten specific subsections measured (e.g., self-criticism, self-satisfaction, physical self, etc.). Yore's findings (1977), however, suggest that efforts to stratify the samples in studies such as those of Udarbe (1973) and Aschermann (1976) may have reduced the sample size to the point where the findings were inconclusive.

Kampsnider's study (1972) of the Lab used in an inservice setting by the Fort Worth (Texas) Independent School District showed evidence of the Lab's influence upon teacher attitude.

While he dealt with different variables from the other studies, it is interesting to note that Kampsnider (1972), the principle author of the Interaction Lab, found a significant impact of sustained duration. It is also of interest to note that students on a number of campuses around the country continue to report their experience in the Interaction Lab as a high point in their program in education, nearly equal to student teaching.

Comparative Studies

Phase IV

Research on the Weber State College program to the present time has been directed toward the improvement of the teacher education program as a whole. In this matter much has been learned and applied, but there is still the constant and ongoing need to evaluate and revise. The launching into Phase IV of the overall plan in which the Weber State College program is compared to other colleges, has finally begun after several years of preliminary discussion, speculation and planning.

The Utah Department of Public Instruction has affiliated with the National Federation of States, whose major goal is to study the effectiveness of teaching. During the past two or three years, meetings and correspondence have been conducted between the Utah State Devision of Teacher Services, the Education Testing Service, and members of the Weber State College teacher education faculty. These preliminary efforts were directed at clarifying how the Weber State College System could be examined to determine its effectiveness.

In the spring of 1978, decisions were made on the nature and direction of the study proposal.

Some Basic Elements of the Proposed Comparative Study Are:

a. It will attempt to determine the teaching skills of a representative number of Weber State College graduates and those from other programs, all of whom are now teaching in Utah schools from Provo on the south to Logan on the north (Wasatch Front).

b. Teachers studied will be selected from those in their first, third and fifth year of teaching, and will largely represent, (in addition to Weber State College graduates), graduates from the three major universities; Utah State University, University of Utah, and Brigham Young University.

c. The record of teacher performance will be obtained basically through use of the APPLE Observation System (Annecdotal Processing to Promote the Learning Experience) used in the California Beginning Teacher Evaluation Study (BTES) and elsewhere. This will be supplemented by use of a "work diary" kept by each teacher, and possibly by other measures, such as supervisor, peer, and pupil ratings.

d. Pupil achievement scores (pre- and post-test) in the classrooms of the teachers studied will be obtained, and categorized into those classes showing high achievement and those showing low achievement.

e. High and low achievement classes will be studied in an attempt to determine those teacher-controlled elements present in the two categories of classes, with a view to identifying cause and effect relations.

f. Teachers at different year levels of teaching will be studied to determine what happens to teaching skills over an extended period of time.

g. Direct comparisons will be made between the results from teachers from different teacher education programs, to determine the prevalence of teaching effectiveness elements and the retention or improvement of skills over an extended period of time.

Possible Hypothesis for Testing:

Weber State College Graduates --

1. Show a high level of initiative, innovativeness, independence,

variety, and creativity.

2. Maintain a positive learning environment through effective use of interpersonal skills.
3. Motivate students to productive activity; students are "on task."
4. Identify student needs and effectively respond to them.
5. Are different in teaching effectiveness from graduates of other teacher education programs.
6. Are capable of meeting the needs of pupils with unusual backgrounds and abilities.
7. Are effective in working with parents and other community members.
8. Are intellectually alert and of a scholarly disposition.
9. Are able to cope effectively with realistic situations.
10. Have a wholesome professional attitude toward teaching.
11. Have a high employment and retention rate.

Duration of Study

It is anticipated the Weber State College teacher effectiveness study will take about three years for the collection and analysis of the data. Reporting the results may extend over several additional years since the multivariate nature of the problem offers almost endless reporting opportunities.

References

Adamson, H. K. An informal survey of schools using the interaction laboratory for teacher development. Unpublished monograph, Weber State College, 1977.

Aschermann, J. R. Human relations training and its effects upon the self-concept, dogmatism and interpersonal attitudes of student teachers. Unpublished doctoral dissertation, Utah State University, 1976.

Boyan, N. J., et al. The instructional supervision training program - Final report: 1972-73. Santa Barbara, California: Graduate School of Education.

Boyan, N. J., & Copeland, W. D. Instructional supervision: Improving the influence of the cooperating teacher. Developing Supervisor Practice, Bulletin 41, pp. 28-38, Association of Teacher Educators.

Burke, C. D. How fare the graduates of a competency-based teacher education system? Journal of the Student Personnel Association for Teacher Education, 13(2), December, 1974.

Burke, C. D., & Stone, D. R. A research-based learning process model for developing and evaluating teacher education curricula. Journal of Teacher Education, 24(3), Fall, 1975.

Kampsnider, J. J. Human relations training for teachers -- A structured laboratory approach. Unpublished doctoral dissertation, University of Utah, 1972.

Low, W. B. (ed.). A guide to student teaching. Weber State College, 1978.

Parker, R. Weber State College evaluates IPTE after three years. Phi Delta Kappen, Bloomington, Indiana, January, 1974.

Udarte, J. R. Interpersonal communication effects on cognitive and affective competencies in an individualized performance-based teacher education program. Unpublished doctoral dissertation, Northern Arizona University, 1973.

Yore, L. D. The effects of the interaction laboratory for teacher development on the self-concept of first year elementary education students. Unpublished monograph, University of Victoria, British Columbia, 1977.

SECONDARY TEACHER PREPARATION PROGRAM
UNIVERSITY OF OREGON
PROCEDURES FOR PROGRAM EVALUATION

Richard I. Arends
University of Oregon

SECONDARY TEACHER PREPARATION PROGRAM
UNIVERSITY OF OREGON
PROCEDURES FOR PROGRAM EVALUATION

Richard I. Arends
University of Oregon

This document reports on the evaluation activities undertaken to assess effectiveness of the Secondary Teacher Preparation Program at the University of Oregon and the competencies of its graduates. It is a summary of the evaluation activities described more fully in the Evaluation Plan for Secondary Teacher Education (Arends and Bullock, 1976); Annual Evaluation Report (Arends, Hesse, Wheeler and Garrett, 1978); and Summary Evaluation Report (Arends and Hesse, 1978).

In August, 1974, a preliminary evaluation plan for the Secondary Teacher Preparation Program at the University of Oregon was initiated by a committee of Secondary Education faculty. This resulted in a survey that was conducted in the spring of 1975, with graduates who had completed the program between 1972 and 1974. Based on the information collected in this survey (see Hesse and Ferry, 1975), several recommendations were made. One recommendation was to extend the original evaluation plan and activities.

Based on what was learned in the 1975 survey and on a review of evaluation plans used at other institutions (e.g., Schalock, 1975; Sandefur, 1975), by the spring of 1976, a comprehensive evaluation plan was subsequently designed and approved by the Secondary Faculty (see Arends and Bullock, 1976). The 1976 Evaluation Plan expanded prior evaluation efforts in two significant ways. First, plans were made to collect information from a larger number of sources: students in the Secondary Program still at the university, graduates of the program at the completion of their first year of teaching, and users of graduates, mainly principals and supervisors in the public schools. Second, the 1976 Evaluation Plan called for the incorporation of evaluation as an ongoing program activity.

Purposes of the Evaluation

Evaluation activities are conducted to collect information so that informed judgments can be made about the effectiveness of the Secondary Program and the competencies of its graduates. The purposes are to guide program development and program review. More specifically, the evaluation submits the program to inspection so that the following outcomes can be assessed:

1. Graduate satisfaction and perceptions of utility and sufficiency of the various components of the Secondary Teacher Preparation Program.
2. Graduate acquisition of the specified competencies of the Secondary Program.
3. Work and career status of University of Oregon graduates.
4. User (educators in the public schools) atttitudes toward the Secondary Program and user perceptions of the competencies of University of Oregon graduates.

Audiences for the Evaluation

Evaluation of the Secondary Program at the University of Oregon is conducted to provide information for the needs of five audiences:

1. Teaching and administrative faculty in the Secondary Education Program and faculty in the Professional Schools at the University of Oregon whose responsibility it is to plan and implement teacher training programs.
2. Students who are considering selecting an undergraduate major in Secondary Education.
3. Public School Educators who are the primary employers of graduates from the Secondary Program.
4. Members of the Consortium for the Improvement of Professional Education (University of Oregon, Eugene; Springfield and Bethel Public Schools) who help plan and provide approval for all University of Oregon certification programs.

5. Members of the Oregon Teacher Standards and Practices Commission (TSPC) and the National Council for Accreditation of Teacher Education (NCATE) who assess the quality and accredit the Secondary Program at the University of Oregon.

Evaluation results are shared with each of the above groups to keep them informed as to the progress of the evaluation activities and to guide future judgments, review and development of the Secondary Program.

Description of the Teacher Education Program Being Evaluated

The University of Oregon adheres to a philosophy of training teachers within a liberal arts environment. Therefore, the preparation of secondary teachers is a joint venture of the various liberal arts departments, the professional schools, and the three divisions of the College of Education. Some students fulfill general studies requirements and prepare for teacher certification within departments of liberal arts: Biology, Chemistry, Physics, Mathematics, English, Romance Languages, German, Russian, Speech, Drama, History, Geography, Anthropology, Economics, Political Science, Psychology, and Sociology. Others are prepared within the various professional schools: the School of Music; the School of Architecture and Allied Arts; the School of Journalism; and the School of Health, Physical Education and Recreation.

The component of the Secondary Program concerned most directly with the theory and practice of teaching and the nature of schooling is the Department of Secondary Education, which is housed within the College of Education's Division of Teacher Education. However, several of the required courses in the professional component are offered by faculty in the Center for Educational Policy and Management and the Division of Developmental Studies and Services. One of the field experiences for secondary students is supervised by ESCAPE, a student-operated unit within the university.

In addition to general studies courses required of all university students, students in secondary education take norm area course work (usually between 45 and 60 credit hours) in one of the liberal arts departments or professional schools. Norm courses in the liberal arts departments usually are not designed especially for teachers; they are part of a curriculum aimed at producing majors in a particular subject field or contributing to a student's liberal education. However, many norm courses in the professional schools are designed especially for students planning careers in teaching or related fields.

Students in the Secondary Program also take 19 credit hours of formal course work in the College of Education. The courses include CI 314, Teaching Strategies; CI 436, Secondary Media; CI 469, Reading and Writing in the Secondary School; EPsy 321, Human Development and Education; EPsy 322, Human Learning and Education; EPsy 323, Psychology and Problems in Education; a special methods class in a student's teaching field; and one of the following courses on social and cultural foundations: EdP 327, Social Foundations of Education; EdP 441, History of Education; EdP 445, Modern Philosophy of Education; or CI 407, Problems of Minorities in Schools. Students do a pre-student-teaching practicum (3 credit hours, supervised by ESCAPE); student teach in a public junior or senior high school for one term (15 credit hours); and take a seminar that accompanies student teaching (1 credit hour).

In addition to course work and field experiences, the university and college provide several support services and resources for students as they prepare for teaching and as they secure their first jobs. These services and resources include the general library, a special curriculum library, the computer center, an audio-visual center, norm area and educational advisors, a field placement office, the certification office, and the Career Planning and Placement Office.

Evaluation Methodology and Activities

The methodology described in the Secondary Evaluation Plan (Arends and Bullock, 1976) identified several areas of the Secondary Program that were to become the focus of the evaluation and identified several evaluation questions to guide the collection of evaluation information. These questions are summarized below:

1. How satisfying, useful and sufficient are various components of the Program:

   as perceived by graduates immediately following student teaching and by graduates one year from student teaching?

as perceived by graduates from liberal arts and professional schools?

2. How competent are graduates of the Program in eleven specified areas:

   as reported by graduates of liberal arts and those who employ or supervise them?

   as reported by graduates of professional schools and those who employ or supervise them?

3. What happens to graduates one year after graduation:

   as reported by graduates from liberal arts?

   as reported by graduates from professional schools?

4. What positive and negative comments and recommendations for revisions of the Secondary Program are made:

   by graduates?

   by employers or supervisors of graduates?

Methods of Obtaining Evaluation Information

To provide answers to the evaluation questions, three interrelated investigations were conducted during the 1976-77 school year: (1) a survey of undergraduates in the 1976-77 Secondary Program immediately following their student teaching experience; (2) a survey of graduates from the 1975-76 Secondary Program one year after their graduation; and (3) an intensive field study of a selected sample of 1975-76 graduates who were teaching in Oregon one year after graduation. Each investigation is described separately below.

Survey of Undergraduates

During the 1976-77 school year, survey information was collected from and about undergraduates in the Secondary Program. The sample for this investigation consisted of 202 undergraduates (62.5 per cent of the total population) who completed student teaching and graduated from the program during fall, winter, or spring terms in 1976-77. Information was collected in the following ways:

Post-Student Teaching Survey Questionnaire. For the fall, winter and spring terms the Post-Student Teaching Survey Questionnaire was administered to all student teachers during the last two weeks of their student teaching experience. This questionnaire aimed at collecting information from students regarding their perceptions of (1) satisfaction, usefulness, sufficiency and recommendations for the Secondary Program; (2) satisfaction, availability and recommendations about program resources; and (3) judgments about specific course work required in the Secondary Program. Questionnaires were mailed directly to students and were returned through self-addressed envelopes or via their university supervisors who had been informed of the survey and instructed to encourage students to complete their questionnaires.

Competency Inventory--Student Version. During spring term, 1977, the Competency Inventory was administered to student teachers during the last week of student teaching. It accompanied the previously described Post-Student Teaching Survey Questionnaire. Students were asked to judge their own competence in eleven areas (specified by the Secondary faculty) which reflected the goals of the program and the competencies needed for successful performance as a teacher. The eleven competence areas included: (1) ability to develop goals and objectives and to plan; (2) ability to select and design curriculum materials; (3) ability to use a variety of teaching strategies; (4) ability to use reading and writing activities; (5) ability to pace and sequence learning activities; (6) ability to use effective classroom management procedures; (7) interpersonal competence and relationships; (8) knowledge of subject matter; (9) ability to evaluate and assess learning; (10) ability to solve problems and innovate; and (11) ability to grow as a professional. For a copy of the Teacher Competency Inventory and other data, write to the author's address as given in the "List of Contributing Authors."

Competencies Inventory--Cooperating Teacher and University Supervisor Version. Also, at the end of spring term, 1977, a Competencies Inventory was administered to the student teachers' cooperating teacher and the university supervisor. These persons were asked to judge the student teachers' level of competence in the same areas included on the student version of the Inventory. These inventories were given directly to the university supervisors to complete, and they were asked to deliver and collect the inventories from the cooperating teachers.

All instruments used in the undergraduate surveys were developed by the evaluation staff of the Secondary Program and were reviewed by the Secondary faculty for appropriateness. Each instrument was pilot-tested and inspected for technical quality.

Survey of First-Year Graduates

During the spring of 1977, names of 1976 graduates were collected from files kept by the Office of Career Planning and Placement at the University of Oregon and from the Certification Offices at the university and at the State Department of Education. Identified graduates were mailed the Follow-Up Survey Questionnaire with detailed instructions and a return address envelope. As with the survey of undergraduates, respondents were asked to make judgements about (1) their satisfaction and perceptions of utility and sufficiency of the Secondary Program, (2) their satisfaction and the availability of program resources and, (3) specific course work required of them when they were going through the program. They were also asked to report on their current work status and plans one year after graduation.

In this investigation, questionnaires were mailed to 169 graduates for whom addresses could be found. Eighty-two questionnaires were returned, for a response rate of 42 percent.

Field Study of First-Year Graduates

From the names collected for the Follow-Up Survey of first year graduates, a stratified sample was selected for more in-depth follow-up and investigation. Twenty-five graduates were selected, using the following criteria: (1) currently teaching in Oregon, (2) indicated willingness to participate in the study and allow an on-site visit, and (3) principal was willing to participate in the study and allow an on-site visit. The sample was selected to include persons from both the liberal arts and the professional schools, and those who were currently teaching in senior high schools and in junior high schools.

Members from the evaluation staff and selected members of the Secondary faculty visited each graduate during May, 1977, and collected information using the following instruments.

Principal Interview. Principals at each site were interviewed and data collected about their perceptions of the first year teachers' competencies and their recommendations for improvement of the Secondary Program.

Teacher Interview. Each first year teacher in the sample was interviewed by the on-site observer. The first year teachers were asked to judge their own level of competence and to make recommendations for the Secondary Program from the perspective of one year of teaching.

Observation. Observers visited each teacher's classroom twice during the on-site visit. Systematic data were collected, and, at the completion of the observation, observers were asked to judge the level of competence of the first year teacher.

Each observer who helped collect information for the Field Study had received three hours of training prior to the on-site visits through the use of a special training packaged created to develop interviewing and observation skills.

Use of Evaluation Results

The information collected in the Evaluation of the Secondary Teacher Preparation Program, to this point, is mainly descriptive. Data are summarized in descriptive tables and charts and are summarized in reports. They are provided to the previously described audiences for the Evaluation.

Three major methods of feedback are employed. First, faculty at the university and members of the University/Public School Consortium are presented with the Annual Evaluation Report (Arends, Hesse, Wheeler and Garrett, 1978). Second, oral presentations are made to explain the information, and the evaluation staff presents the faculty with recommendations based on the evaluation information.

Finally, a special report (Summary Evaluation Report, Arrends and Hesse, 1978) is prepared for and mailed to first year teachers and their principals wbo participated in the evaluation. The report is also given to others, such as prospective students and employers of university graduates or to those who ask for evaluation information. Data in these summary reports are presented in graphic form for clarity of understanding and ease of interpretation.

Strengths, Weaknesses and Future Plans

The strengths of the present evaluation activities used at the University of Oregon for evaluation of the Secondary Teacher Preparation Programs can be summarized in the following statements:

1. A large amount of information about the programs and its graduates that is potentially useful for program development is collected at a relatively small cost.
2. Using faculty and members of the University/Public School Consortium to plan the evaluation and collect some of the evaluation information increases the likelihood that it will be used. This involvement has provided a self-reflective attitude by the faculty toward program development and a less critical, more cooperative stance by public school personnel who have been involved.

At the same time, there are several weaknesses and inadequacies of present efforts. For example:

1. The information currently being collected relies primarily on high inference judgments by untrained raters and self-report by graduates and first year teachers.
2. No resources have been allocated that would allow looking at the relationship between any number of independent variables and the overall competence or effectiveness of graduates.
3. Likewise, no resources have been allocated to obtain measures of graduate effectiveness as portrayed by student achievement, classroom climate, etc.
4. Even though the field study provides richer information about graduates as compared to a survey, the evaluation does not provide the rich, illustrative information about graduates that could be acquired through intensive case studies and analyses.

Given present resource allocations for providing evaluation of the Secondary Program, plans for the future call for continuation of present efforts but no extensions. Instruments used in the evaluation will continue to be refined and procedures for collecting and storing information simplified.

It is not possible at the present time to collect the type of information that would correct many of the inadequacies described above.

References

Arends, R. I., & Bullock, T. Evaluation plan for secondary teacher education. Eugene, Oregon: University of Oregon, 1976.

Arends, R. I., Hesse, K, Wheeler, S., & Garrett, J. Annual evaluation report. Eugene, Oregon: University of Oregon, 1978.

Arends, R. I., & Hesse, K. Summary evaluation report. Eugene, Oregon: University of Oregon, 1978.

Hesse, K., & Ferry, C. Follow-up survey of first-year graduates between 1972 and 1974. Eugene, Oregon: University of Oregon, 1975.

Sandefur, J. T. Western Kentucky University teacher preparation evaluation system. Bowling Green, Kentucky: Western Kentucky University, 1974.

Schalock, D., et al. Plan procedures for study of graduates. Monmouth, Oregon: Teaching Research, 1976.

THE OHIO STATE UNIVERSITY
FOLLOW-UP STUDY

Judith D. Aubrecht, Kevin Ryan, Mikell O'Donnell
The Ohio State University

THE OHIO STATE UNIVERSITY FOLLOW-UP STUDY

Judith Aubrecht, Kevin Ryan, and Mikell O'Donnell
The Ohio State University

Problem Statement

The College of Education of The Ohio State University has twenty-five undergraduate teacher certification programs. The intent of the Follow-up study project conducted by the Office of Program Development and the Office of Student Development with the assistance of the College of Education was to design an evaluation system which would elicit information from and about graduates of these twenty-five programs. The information collected needed to be specific enough to be useful to individual program groups and general enough to provide both a general college overview and comparisons among programs.

Key References

Follow-up studies seem to fall into two general categories. One could be labeled "teacher effectiveness" and the other "program evaluation." In the first, a model of "good teaching" is explicitly set forth. The extent to which program graduates fit the model is then tested through some combination of survey, observations, and personality inventories (Sandefur, 1972; Sandefur and Adams, 1976; Adams, 1974). In the program evaluation category, the elements of an existing teacher education program (courses, services, experiences) are described, and graduates are asked to rate the quality of these elements (Redwine, 1973; Ayers, 1976).

Data gathering techniques in follow-up studies have typically included some combination of the following: (1) questionnaires on demographic data, (2) survey instruments on teacher behaviors, (3) survey instruments on program elements, (4) systematic observation (usually a Flanders-type instrument), (5) supervisor evaluation forms, and (6) personality inventories. Sometimes an attempt is made to correlate the data obtained through different techniques. Student evaluations may be correlated with personality inventories; teacher behavior (from observation) may be correlated with either student evaluations or personality inventories (Sandefur and Adams, 1976).

Few teacher education programs have reported extensive follow-up studies. It is safe to assume that the situation will change rapidly in the next few years. It is also probably safe to assume that most of the work in the near future will be based on ideas of teacher effectiveness. Even if the goal is program evaluation, the emphasis will be on the behavior of the teacher in the field. Changes in teacher education programs will come largely through the assessment of the needs and abilities of program graduates.

Most of the research on teacher effectiveness has been conducted at the university level and most of it has dealt with student evaluations and their potential impact on teacher behavior (Coley, 1975; Menges, 1973). Other attempts to evaluate and/or enhance teacher effectiveness through videotaping, observation, student interview, and teacher self-ratings are in their infancy. For this reason, the literature on teacher effectiveness is helpful to us only in general ways. Our purpose is not to induce changes in the teaching behavior of graduates of the programs. Nor is it really to evaluate these graduates as teachers (although this is inevitably a component of the process). It is rather to assess their needs and abilities, and, using the results of this assessment, to determine the implications for our programs. We learn from the literature that it is useful to provide a wide variety of ways to study these graduates. Among the ways are: (1) survey questionnaires for graduates, (2) interviews with graduates, (3) questionnaires for students of these graduates, (4) questionnaires for supervisors and peers (other teachers), and (5) systematic observation data from the graduates' classrooms.

Rationale Objectives

The purpose of any evaluation system is to collect, analyze, and disseminate information that is useful for decision-making purposes. The primary need in a follow-up evaluation system for an institution as complex as The Ohio State University (OSU) College of Education is to generate information that will serve the decision-making needs of individual program groups. In addition, there are individuals and groups that can be served by selected portions of the data or by various kinds of summary data generated within the system. These include faculty teaching "core" courses (those courses that are taken by students in all programs), faculty directing field experiences shared across programs, curriculum redesign groups, Student Services, and the Dean's staff.

Certain basic assumptions have guided the design of the evaluation system at OSU.

1. In order to assure that the system would generate useful and persuasive data, faculty involvement in the design stage was imperative. In the early weeks of the project, the intent was to have project staff work with each program group to generate instruments reflecting individual program needs. It soon became apparent that, with the time allocated for project completion, this strategy was not feasible. Instead, it was decided that program groups would be asked to participate in generating instruments and that review committees would meet periodically to advise project staff.

2. Statements of the program's "goals and objectives" were considered too general to be useful as the basis for instrument items. Statements of teacher "competencies" were considered too specific to be useful in this way. "Observable teacher behaviors" provided a middle level of generality and therefore, were selected as the basis for instrument design. Teacher attitudes, while not necessarily observable, were also considered important factors for assessment.

3. Faculty in twenty-five diverse program groups were asked to specify "desirable teacher behaviors and attitudes." They were to report: (a) some "core" set of these behaviors and attitudes, and (b) some sets of behaviors and attitudes that were program specific. A "core" survey instrument would be based on the consensus across programs (plus contributions from "core" courses and experiences and contributions from curriculum redesign groups). The individual program survey instrument would be based on the program-specific items identified and agreed upon by faculty within program groups.

4. Demographic and "school climate" information would be important in interpreting survey and observation data.

5. The "useful and persuasive" information would appear in the comparisons of ratings of "desirable attitudes and behaviors" by program faculty, program graduates, supervisors of program graduates, peers of program graduates, and students of program graduates. Further information

> would appear in comparisons between the students' ratings of "desirable attitudes and behaviors" and their own perceptions of their abilities to behave in ways that they value. Corresponding values and perceptions of teacher ability by supervisors and peers, and student perception of teacher performance would provide a reality check on graduates' responses. The picture would be rounded out by interviews of graduates and by direct classroom observation.

Methodology of the Study

Project staff assembled a pool of 119 items representing "desirable teacher attitudes and behaviors." These items were culled from the literature, from Ohio State Standards, and from staff brainstorming. Items were grouped under rational headings. Some of these headings (Expositional Skills, Definition of Student Responsibilities, Relevancy of Course Material, Stimulation of Ideas and Thinking, Tolerance of Other Viewpoints, Attitudes Toward Students) were derived from work by Whitely and Doyle (1976). Other headings (Evaluation of Student Work, Working with Exceptional Children, and Educational Media) were added by project staff. Later in the project, two more categories were added, one on Evaluation (somewhat more general than "Evaluation of Student Work" and including teacher self-assessment) and one on Multi-Cultural Understanding. It was thought that these category headings would help program groups organize their contributions to instrument design.

The pool of items was forwarded to program groups, with a request for their faculty to come to a consensus on items that should be included in a "core" instrument and items that should be included in an individual program instrument. Faculty were encouraged to write their own items, especially those that would be program-specific. During the same period, faculty involved with "core" courses and experiences and the existing curriculum redesign groups were asked to supply items for the "core" surveys.

Two separate "core" surveys were assembled from the faculty consensus materials, one for elementary programs and one for secondary programs. The items selected for the surveys were those most frequently chosen by program groups and by the other contributing groups. Once the "core" surveys were assembled, individual program instruments were composed of the additional items selected by program groups. An attempt was made to eliminate overlapping items (no item appears in both the "core" and individual surveys).

The procedure involved in selecting observation instruments was considerably simpler. A set of instruments was assembled by project staff and reviewed by a committee of faculty advisors and program directors. The "core" set of instruments that emerged included three sets of categories from the Language of the Classroom system developed by Arno Bellack, et al. (1966). These categories were SPEAKER (teacher, pupil) and PEDAGOGICAL MOVE (structuring, soliciting, responding, reacting) and SUBSTANTIVE - LOGICAL MEANING (defining, interpreting, fact stating, explaining, opining, justifying). Also selected were some informal techniques developed by Acheson and Hansen (1973), including Selective Verbatim (teacher questions, student questions, teacher reactions to student talk, teacher "controlling" statements) and an At Task Analysis.

Figure I, PROJECT FLOW CHART/Design Phase, gives an overview of project activity just described. Figure II, PROJECT TIMELINE, contains a more complete task breakdown, along with indications of actual project time involved during the period of January to June, 1977.

Several tasks remained before the pilot stage of the project. Program groups were given the opportunity to examine the "core" surveys in final form and to revise their individual program surveys to reflect more specific program needs. (Later, during the first implementation, Likert scales were attached to the revised program instruments and to the "core" surveys and these were returned to faculty members as a form of content validation. See Figure III, SAMPLE OF FACULTY SURVEY).

Each year a pre-implementation procedure is required in order to select two or three program groups for on-site observation. Program-specific observation techniques are developed and added to the "core" techniques. Observers are trained and inter-observer reliability is established. During the first implementation (1977-78) Early and Middle Childhood Education and Physical Education were the groups selected for observation.

Another pre-implementation procedure to be followed each year is the selection of graduates for the study. This is done through a stratified sampling design. All graduates of Teacher Education programs are surveyed one year out of the program, three years out of the program, and five years out of the program (starting with 1977 graduates). For those two or three program groups being observed in any given year, all graduates are surveyed (by mail) and those teaching within a fifty mile radius of Columbus are also observed (to a maximum sample of thirty). For those graduates who are observed, questionnaire data is also collected from supervisors, peers and client students. At least two follow-up mailings are sent to non-respondent graduates. On-site visits (including classroom observation, interviews of graduates, and surveys of supervisors, peers and students) take place during the same quarter that surveys are mailed.

Each year the survey and observation data is assembled, analyzed and disseminated to the appropriate groups. Desired revisions in instrumentation are indicated by the groups served. Revisions are made prior to the next round of data collection.

Figure IV, PROJECT FLOW CHART/Pilot and Pre-implementation Phase and PROJECT FLOW CHART/Implementation Phase, presents an overview of the activities described above. Figure V presents in a facet design a theoretical overview of the entire follow-up system.

During the first year of implementation (1977-78) many procedures were developed that would be used in following years. Special tasks included the design of optical scanning sheets for surveys; the development of a computer system to read the scanning sheets, analyze the data, and produce reports for various program groups; and the system for obtaining names and addresses for program graduates (as well as mailing procedures). Figure VI, TASK ANALYSIS - FOLLOW-UP IMPLEMENTATION, details the activities of the first year of implementation.

Analysis: Reliability

Survey instruments were pilot tested on groups of supervisors, teachers, and public school children. Instruments were checked for clarity of directions and individual item clarity. Efforts were made to establish uniformity in the administration of student surveys. Surveys were factor analyzed and individual items were correlated with the overall factor group. (See Nunnally, 1967).

For the observation instruments inter-observer reliability was established using Scott's Pi (1955). Interview procedures were checked for clarity of directions, uniformity of administration, and individual question clarity.

Analysis: Validity

The issue of content validity was addressed mainly through faculty selection of items by consensus (during the design phase). A further check could be made after the first implementation when the faculty used the Likert scales to rate the items.

Concurrent validity could be addressed after implementation by determining the degree of correspondence between graduates' responses to surveys and the responses of supervisors and peers. Observation data could also be compared with the self-assessments of the graduates.

The validity of the observation instruments selected had been established through use in prior studies. The content validity of interview schedules was addressed in a limited way during the pilot stage by additional questions to interviewees that probed the meaning of interview questions.

Plans for Data Analysis

The computer system to analyze survey data was developed by Nancy Lee, Al Stutz, and Aaron Supowit of the Instruction and Research Computer Center at OSU. The system generates reports on individual graduates within the programs selected for observation, displaying data from graduate, supervisor, peer and students in close proximity. It summarizes responses within program groups by reporting medians and standard deviations. It also provides summary reports across all elementary and all secondary programs. The system is set up to test associations among survey items and demographic or school climate variables.

Individual program groups receive the report of their own graduates' responses as well as a summary report across programs for comparison. Those program groups selected for observation receive the additional data collected from supervisors, peers, students and through interview and classroom observation. Summary survey data is provided to faculty teaching "core" courses, to faculty supervising field experiences shared across programs, to curriculum redesign groups, to Student Services, and to the Dean's staff.

Implications

The follow-up system that has been presented is inductive and eclectic. Implied in the design is a rejection of any single "ideal teacher" model. At the same time, the search for faculty consensus does imply a belief in the existence of some characteristic behaviors and attitudes that are associated with effective teaching. It is not implied that an exhaustive list of these characteristics has been developed. Nor, of course, is it likely that any one teacher would possess all the characteristics identified. This approach does allow both for general ("core") characteristics that emerge across varied subject fields at the secondary level and for some very specific characteristics that might distinguish the science teacher from the social studies teacher or the dance teacher from the home economics teacher. Although separate "core" surveys were developed for elementary and secondary teachers, the final instruments were strikingly similar. The design accomodates both the idea that a teacher is a teacher no matter what the level or subject matter and the idea that a music teacher is and should be distinguishable from a health teacher in terms of desirable teaching behaviors and attitudes.

If this evaluation system is successful there should be some implications for action within the college. Among these would be:

1. Integrated revisions in college-wide "core" program and individual programs based on the needs and abilities of graduates.

2. Development of inservice programs to serve the needs of practicing teachers.

3. Changes in admission procedures and student services based on a combination of existing data bases and the information yielded by the follow-up system.

4. The evolution of administrative procedures designed to support and encourage the kinds of changes described above.

References

Acheson, K. A., & Hanse, J. H. Classroom observations and conferences with teachers. Burlingame, California: Association of California School Administrators, 1973.

Adams, R. Follow-up and evaluation of teacher education graduates - A pilot study, final report. Bowling Green, Kentucky; Western Kentucky University, Office of Educational Research, 1974.

Ayers, J. B. Tennessee Technological University teacher evaluation model - Year III. Cookeville, Tennessee: Tennessee Technological University, College of Education, 1976.

Bellack, A. A., Kliebard, F., Hyman, R., & Smith, F. T. Language of the classroom. New York: Teachers College Press, 1966.

Coley, R. J. Student evaluation of teacher effectiveness, TM Report 52. Princeton, New Jersey, September, 1975. (ERIC Document Reproduction Service No. ED 117194)

Menges, R. J. The new reporters: Students rate instruction. In C. R. Pace (ed.), New directions in higher education: Evaluating learning and teaching. San Francisco: Jossey-Bass, 1973.

Nunnally, J. C. Psychometric theory. New York: McGraw-Hill, 1967.

Redwine, J. A follow-up of 1967-72 division of education graduates of Indiana University at South Bend. Denver: Educational Resources Information Center, 1973.

Sandefur, J. An illustrated model for the evaluation of teacher education graduates. Washington: American Association of Colleges for Teacher Education, 1972.

Sandefur, J., & Adams, R. An evaluation of teaching: An interim research report. Journal of Teacher Education, 27(1), Spring, 1976.

Scott, W. A. Reliability of content analysis: A case of nominal coding. The Public Opinion Quarterly, 19, 1955, 321-325.

Whitely, S. E., & Doyle, K. O. Implicit Theories in Student Ratings. American Educational Research Journal, 13(4), Fall, 1976.

PROJECT FLOW CHART
Design Phase

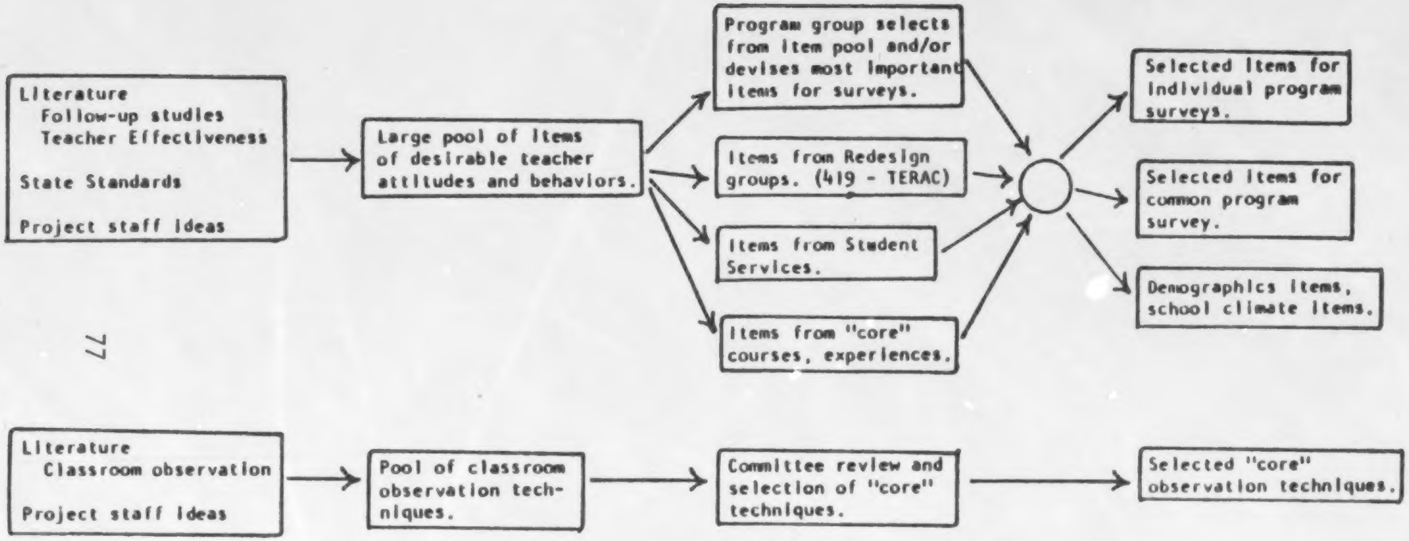
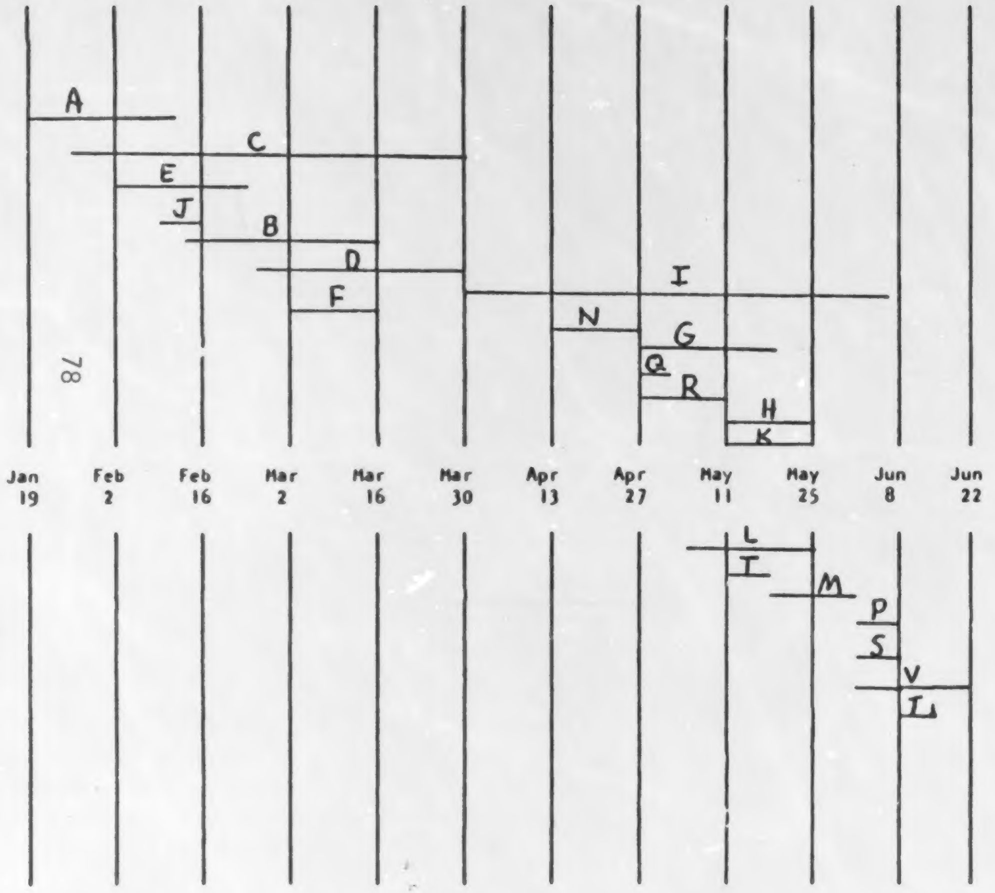

Literature
Follow-up studies
Teacher Effectiveness
State Standards
Project staff ideas
Large pool of items of desirable teacher attitudes and behaviors.
Program group selects from item pool and/or devises most important items for surveys.
Items from Redesign groups. (419 - TERAC)
Items from Student Services.
Items from "core" courses, experiences.
Selected items for individual program surveys.
Selected items for common program survey.
Demographics items, school climate items.
Literature
Classroom observation
Project staff ideas
Pool of classroom observation techniques.
Committee review and selection of "core" techniques.
Selected "core" observation techniques.

Figure I

PROJECT TIMELINE



A. Problem definition. Production of schematic representations.
B. Development of item pool of desirable teacher attitudes and behaviors.
C. Review of the literature on follow-up studies.
D. Review of the literature on teacher effectiveness.
E. Study scaling techniques for surveys. Select appropriate techniques. Determine methods of analysis.
F. Incorporate State Standards into item pool.
G. Incorporate items from student services.
H. Incorporate items from "core" courses, experiences.
I. Have program groups select important items from pool (or write items) for their own program and for "core."
J. Assemble examples of observation tools.
K. Incorporate items from redesign groups.
L. Complete "core" survey.
M. Complete individual program surveys.
N. Design "common" observation techniques.
O. Design observation techniques for individual programs.*
P. Design questionnaires for supervisors, peers, client students.
Q. Submit observation instruments for review.
R. Select appropriate data analysis techniques.
S. Submit surveys for review.
T. Revise instruments.
U. Prepare data reporting forms.*
V. Prepare final report to state.
W. Prepare policy and procedures manual.*

* These steps will be taken during implementation.

Figure II

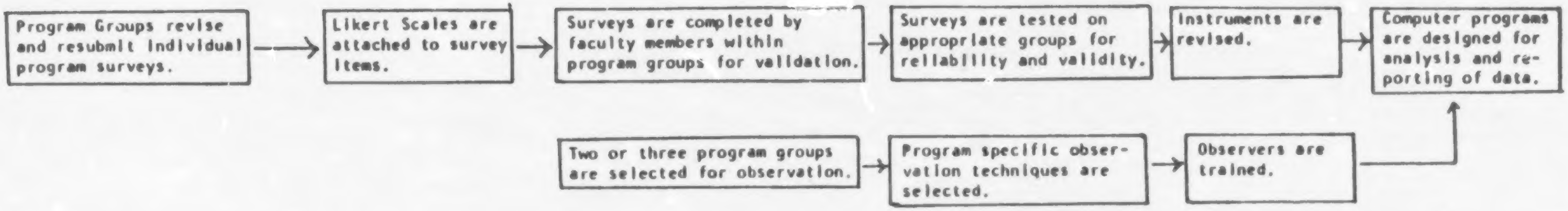
FOR SURVEY OF FACULTY: SAMPLE

Each of the items below describes some behavior or attitude about teaching. Please indicate how important you feel it would be for graduates of your program to display these behaviors or attitudes.

		ESSENTIAL	VERY IMPORTANT	IMPORTANT	DESIRABLE	UNDESIRABLE
A 1.	The teacher presents material at an appropriate level for students.	___	___	___	___	___
A2.	The teacher brings a wide range of instructional methods to bear on a particular learning situation.	___	___	___	___	___
A 3.	The teacher's lessons are carefully planned and have definite purpose.	___	___	___	___	___
A 4.	The teacher maintains progress toward course objectives.	___	___	___	___	___
A 5.	The teacher organizes, interprets, explains and illustrates concepts and relationships.	___	___	___	___	___
A 6.	The teacher's command of the language is adequate.	___	___	___	___	___
A 7.	The teacher communicates subject matter well.	___	___	___	___	___

Figure III

PROJECT FLOW CHART
Pilot and Pre-Implementation Phase


Program Groups revise and resubmit individual program surveys.
Likert Scales are attached to survey items.
Surveys are completed by faculty members within program groups for validation.
Surveys are tested on appropriate groups for reliability and validity.
Instruments are revised.
Computer programs are designed for analysis and reporting of data.
Two or three program groups are selected for observation.
Program specific observation techniques are selected.
Observers are trained.

PROJECT FLOW CHART
Implementation Phase

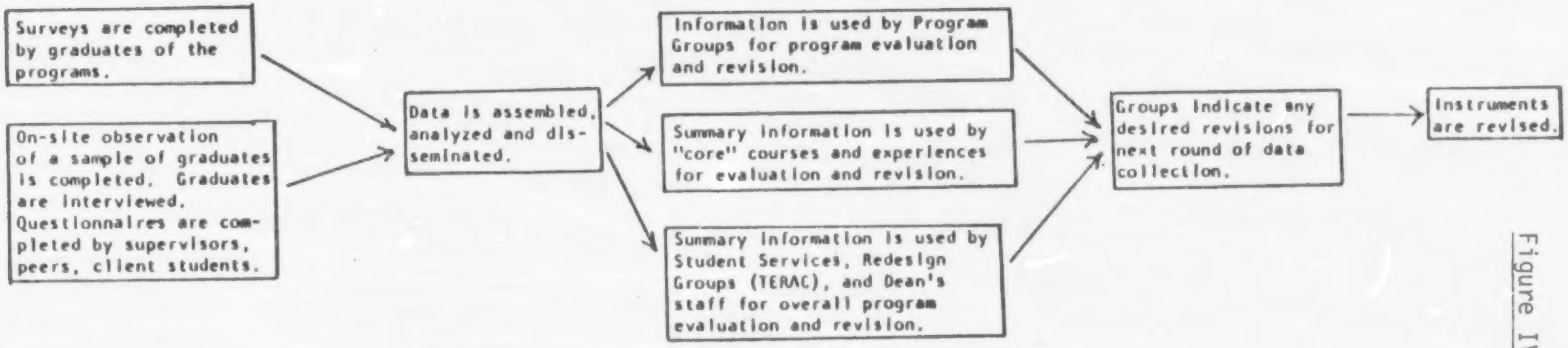

Surveys are completed by graduates of the programs.
On-site observation of a sample of graduates is completed. Graduates are interviewed. Questionnaires are completed by supervisors, peers, client students.
Data is assembled, analyzed and disseminated.
Information is used by Program Groups for program evaluation and revision.
Summary information is used by "core" courses and experiences for evaluation and revision.
Summary information is used by Student Services, Redesign Groups (TERAC), and Dean's staff for overall program evaluation and revision.
Groups indicate any desired revisions for next round of data collection.
Instruments are revised.

Figure IV

The behavior $\begin{pmatrix} x_1 & \text{attitudes} \\ x_2 & \text{actions} \end{pmatrix}$ of the population $\begin{pmatrix} y_1 & \text{teachers from program 1} \\ y_2 & \text{teachers from program 2} \\ \vdots & \\ y_n & \text{teachers from program n} \end{pmatrix}$ towards the program (* see note below) $\begin{pmatrix} z_e & \text{common aspects of elementary programs} \\ z_s & \text{common aspects of secondary programs} \\ z_1 & \text{aspects of program 1} \\ \vdots & \\ z_n & \text{aspects of program n} \end{pmatrix}$

as perceived or demonstrated by $\begin{pmatrix} \text{teachers from the population defined} \\ \text{supervisors of those teachers} \\ \text{peers of those teachers} \\ \text{students of those teachers} \end{pmatrix}$ sampled through $\begin{pmatrix} \text{survey} \\ \text{observation} \\ \text{interview} \end{pmatrix}$ at a time $\begin{pmatrix} \text{after student teaching} \\ \text{one year after program completion} \\ \text{three years after program completion} \\ \text{five years after program completion} \end{pmatrix}$

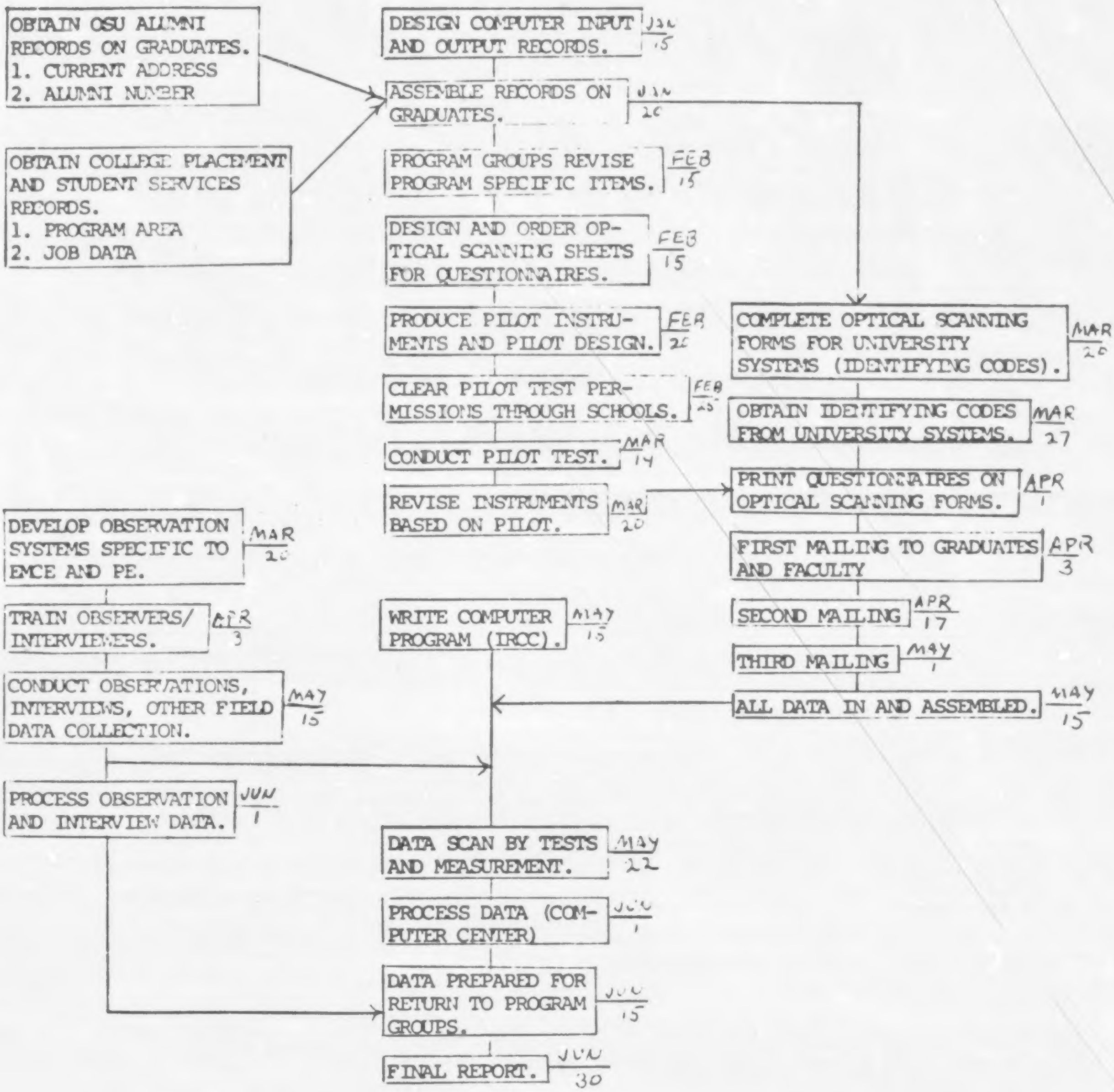
yields a range of information to be used by $\begin{pmatrix} \text{Vice Provost for Instruction} \\ \text{Dean's staff} \\ \text{Redesign groups} \\ \text{Program Directors} \\ \text{Program Committees of the Faculties} \\ \text{Faculty Chairpersons} \end{pmatrix}$ for purposes of program evaluation and change.

* Aspects of programs defined by or derived from:
Program Groups
State Standards
Student Services
"core" courses, experiences
literature, other follow-up studies
redesign groups (TERAC)

Figure V

Figure VI

TASK ANALYSIS - FOLLOW-UP IMPLEMENTATION


OBTAIN OSU ALUMNI RECORDS ON GRADUATES.
1. CURRENT ADDRESS
2. ALUMNI NUMBER
OBTAIN COLLEGE PLACEMENT AND STUDENT SERVICES RECORDS.
1. PROGRAM AREA
2. JOB DATA
DESIGN COMPUTER INPUT AND OUTPUT RECORDS. JAN 15
ASSEMBLE RECORDS ON GRADUATES. JAN 20
PROGRAM GROUPS REVISE PROGRAM SPECIFIC ITEMS. FEB 15
DESIGN AND ORDER OPTICAL SCANNING SHEETS FOR QUESTIONNAIRES. FEB 15
PRODUCE PILOT INSTRUMENTS AND PILOT DESIGN. FEB 20
CLEAR PILOT TEST PERMISSIONS THROUGH SCHOOLS. FEB 25
CONDUCT PILOT TEST. MAR 14
REVISE INSTRUMENTS BASED ON PILOT. MAR 20
COMPLETE OPTICAL SCANNING FORMS FOR UNIVERSITY SYSTEMS (IDENTIFYING CODES). MAR 20
OBTAIN IDENTIFYING CODES FROM UNIVERSITY SYSTEMS. MAR 27
PRINT QUESTIONNAIRES ON OPTICAL SCANNING FORMS. APR 1
FIRST MAILING TO GRADUATES AND FACULTY APR 3
SECOND MAILING APR 17
THIRD MAILING MAY 1
ALL DATA IN AND ASSEMBLED. MAY 15
DEVELOP OBSERVATION SYSTEMS SPECIFIC TO EMCE AND PE. MAR 20
TRAIN OBSERVERS/ INTERVIEWERS. APR 3
CONDUCT OBSERVATIONS, INTERVIEWS, OTHER FIELD DATA COLLECTION. MAY 15
PROCESS OBSERVATION AND INTERVIEW DATA. JUN 1
WRITE COMPUTER PROGRAM (IRCC). MAY 15
DATA SCAN BY TESTS AND MEASUREMENT. MAY 22
PROCESS DATA (COMPUTER CENTER) JUN 1
DATA PREPARED FOR RETURN TO PROGRAM GROUPS. JUN 15
FINAL REPORT. JUN 30

FOLLOW-UP STUDIES IN SCIENCE EDUCATION AT THE OHIO STATE UNIVERSITY

Patricia E. Blosser, Robert W. Howe
The Ohio State University

FOLLOW-UP STUDIES IN SCIENCE EDUCATION
AT THE OHIO STATE UNIVERSITY

Patricia E. Blosser

and

Robert W. Howe
The Ohio State University

Overview Description

The improvement of the teacher education program in the College of Education at The Ohio State University has been the concern of many faculty members and graduate students. To address this concern, since 1964, about 7 faculty members and 15 graduate students in science and mathematics education have been involved in follow-up activities related to the preservice program for secondary school science and mathematics teachers.

Seven doctoral dissertations, focused on the science education portion of the program, and two dissertations, focused on the mathematics education portion of the program, have resulted from these follow-up activities. In addition, some informal research activities are being carried on that have not yet resulted in formal data analyses and publications. For the purposes of this report, discussion will center on the work completed relative to the science education program and activities, although the dissertations in mathematics education will be identified and described briefly.

The initial purposes of the follow-up studies were:

1. To determine strengths and weaknesses of the science education program, based on evaluation of teacher performance.
    a. To determine what program experiences graduates found to be most useful in their teaching activities.
    b. To determine what program experiences students found least useful in their teaching activities.
2. To determine the style(s) of teaching used by recent graduates.

3. To determine the possible effects the school setting had on each teacher's success.
4. To determine the attitudes of the inservice teacher and to compare these with the attitudes the teacher had as a preservice teacher.
5. To determine the apparent success of the teacher and compare this to the success expected by the college staff.
6. To consider program changes based on the results of the data analyses.

While a variety of designs and methodologies were used in the studies, several of them used similar approaches. Data were normally collected on inservice teachers (our graduates) both when they were enrolled in the preservice program and during the early years (1-5) of their teaching career. Since 1967, the studies have been limited to a study of our graduates who remain in Ohio.

History of Follow-Up Studies at Ohio State University

Follow-up studies have been conducted at Ohio State University since 1964. 1975-76 was the last year in which a study of the teacher education program in science and mathematics at Ohio State University (OSU) was completed. Further activities have been delayed until the College of Education determines a future course of action for follow-up studies of its graduates. Several options are being reviewed by the College regarding ways research in science and mathematics education might continue.

In the late 1960's the preservice teacher preparation program in science and mathematics education at OSU was modified to incorporate extended field experiences prior to student teaching and to involve preservice students in a variety of educational levels (elementary, junior high school, senior high school) and community settings (inner city, suburban). Follow-up studies focused on determining the impact of these experiences on the preservice teachers and on identifying changes, if any, in the behaviors of these students as they became inservice teachers in science education (Sagness, 1970; Brewington, 1971; Cignetti, 1971; Brown, 1972). Two studies similar in intent to those in science education were completed in mathematics education

(Graening, 1971; Erb, 1971).

Additional facets of the science education program were studied as well. Lucy (1972) examined the effect of the laboratory activities component of the preservice program on teacher behavior. Deamer (1973) attempted to obtain information concerning the perceptions of preservice students (prior to and during student teaching), college supervisors, cooperating teachers, and administrators regarding the value of the various objectives of each of the five quarters of the preservice program. Swami (1975) followed up graduates, one to five years after completion of the program at OSU, to determine their feelings about and their use of inquiry-oriented science activities in classrooms.

Measures (Instruments) and Data Base

Data were collected on OSU students both while they were enrolled in the preservice program and after graduation, when they were inservice teachers (with 1-5 years of experience). Date were also collected from the pupils of graduates in a selected class (or classes) and from administrators or supervisors in the school in which the graduate taught.

A number of studies also involved analyses of the preservice students by science education faculty members. In other cases, informal analyses of data available regarding ratings of students by their college supervisor and cooperating teachers were used.

After the school was identified in which one or more of the graduates was teaching, permission was obtained from the graduates and from the principal in each school to conduct follow-up research. We have usually obtained good cooperation from most people. A few teachers, usually those who have had problems with our program, have elected not to participate. Data were collected from teachers and their pupils and from administrators. Questionnaires, inventories, and usually interviews, were used.

An attempt was also made to collect data on the school setting. Normally this was done was using data supplied by the teacher (our graduate), by the principal or science supervisor, and by the OSU staff. The type of data collected has varied, depending on a number of variables.

When data were available on the teacher (our graduate) as a preservice student, a pre-post analysis was done. Other data analyses were primarily post. Teacher style, ratings of teacher effectiveness, and attitudes of teachers were used for analysis.

While instruments varied across studies, several were consistently used. These are the Science Classroom Activities Checklist: Teacher Perceptions (SCACL:TP); Science Classroom Activities Checklist: Student Perceptions (SCACL:SP); Checklist for Assessment of Science Teachers: Supervisor's Perceptions (CAST:SP); Checklist for Assessment of Science Teachers: Pupil Perceptions (CAST:PP). The two versions of the SCACL (teacher and student) were developed by Sagness (1970).

The SCACL:TP and SCACL:SP instruments have been influenced by research done by Kochendorfer and Lee. Kochendorfer was interested in determining the degree to which a teacher's classroom practices agreed with those practices thought to contribute positively toward the attainment of BSCS (Biological Sciences Curriculum Study) objectives. Sagness modified and rewrote the instrument so that it was applicable regardless of the science discipline involved. The SCACL instruments are designed to determine the nature of classroom activities which teachers feel should be used for secondary school science instruction. The Student Perception version of the SCACL serves as a check on how students perceive their teachers actually behaving. The SCACL:TP and SCACL:SP relate to teaching style in terms of use of inquiry-oriented science activities.

The SCACL contains seven subscales: (1) Student Classroom Participation, (2) Role of the Teacher in the Classroom, (3) Use of Textbook and Reference Materials, (4) Design and Use of Tests, (5) Laboratory Preparation, (6) Type of Laboratory Activities, and (7) Laboratory Follow-Up Activities.

In his 1972 dissertation, Brown reports (Table 17, page 75) Kuder-Richardson 20 and 21 reliabilities for the revised SCACL. Brown's information is reproduced in Table 1.

The Checklist for Assessment of Science Teachers (CAST) was developed in two forms in order to assess characteristics of science teachers (Brown, 1972). CAST items focus on three major areas: (1) student-teacher relations, (2) classroom activities, and (3) teacher's personal adjustment. The pupil

Table 1*

KUDER-RICHARDSON 20 AND 21 RELIABILITIES
FOR THE REVISED SCIENCE CLASSROOM
ACTIVITY CHECKLIST

	1969-70		1970-71	
Revised SCACL	KR-20	KR-21	KR-20	KR-21
Teacher's Perceptions-Urban	.75	.70	x	x
Teacher's Perceptions-Suburban	.80	.76	x	x
Student's Perceptions on the Cooperating Teachers	.77	.73	.71	.67
Student's Perceptions on the Student Teachers	.74	.71	.71	.67

*From Brown (1972) Table 17, p. 75

perceptions form (CAST:PP) assesses only student-teacher relations and classroom activities. The supervisor's perceptions form (CAST:SP) assesses all three areas.

The reliability of the CAST was computed by the use of The New Item Analysis Program developed by the Office of Evaluation at The Ohio State University. A KR-20 of .74 and a KR-21 of .71 were obtained for the CAST:PP with 327 students. Validity of the CAST was also computed, using ratings from doctoral students and professors in science education at The Ohio State University. Table 2 reproduces the results of Brown's 1972 study (Brown, Table 18, page 82).

Data Base Descriptions

Data obtained in the various studies were coded for key punching and computer analysis. The types of information contained in the data base are best exemplified by the coding information provided on the 124 variables in Swami's study (1975).

Table 2*

INTRACLASS CORRELATION AND INTRACLASS CORRELATION OF THE SUM OF THE RATINGS FOR THE CHECKLIST FOR ASSESSMENT OF SCIENCE TEACHING

	Questions	r_{cc}[a]	r_{kk}
1.	Teacher's disciplinary ability	.78	.98
2.	Student or subject matter point of view	.94	.99
3.	Teacher's attitude toward adolescents	.86	.99
4.	Teacher understand behavior problems	.76	.97
5.	Attitude of students toward teacher	.96	.99
6.	Student's role in class	.98	.99
7.	Teacher's role in class	.80	.94
8.	Use of textbook and reference materials	.88	.99
9.	Design and use of tests	.89	.99
10.	Conduction of laboratory	.93	.99
11.	Teacher analytical thinking	.53	.93
12.	Teacher social attitudes	.98	.99
13.	Teacher emotional attitudes	.97	.98
14.	Teacher self-confidence	.86	.99
15.	Teacher personal relations	.86	.99

[a] Intraclass Correlation:

$$r_{cc} = \frac{V_r - V_e}{V_r + (k-1)\,V_e}$$

Intraclass Correlation of the Sum of the Ratings:

$$r_{kk} = \frac{V_r - V_c}{V_r}$$

V_r = variance between rows, where each row stands for a person (ratee)

V_e = variance for residuals or error

k = number of columns (raters)

*From Brown (1972) Table 18, p. 82

In addition, computer card decks were available from some doctoral dissertations. One or two students put their data on tape, but this has not been routinely done. Much of the data were reported in tables and appendices in the various dissertations.

In addition, there were informal sources of data in student files. Students are evaluated during each of the five quarters they are enrolled in the preservice program. Cooperating teachers are routinely asked to evaluate the students, as are the college supervisors. In addition, each student is expected to complete a self-evaluation each quarter. Data from these evaluations vary from year to year, depending on the faculty member(s) in charge of a particular portion of the preservice program.

Feedback Procedures

Information from doctoral dissertations has been used for modification of the teacher education programs in science and mathematics at OSU. Feedback from cooperating teachers and administrators, as well as feedback from students who have completed the preservice program in science and mathematics education, has been used in seeking program support from the College of Education administration.

In addition, the more informal kinds of data gathering activities identified earlier in this paper have been used for counseling purposes with individual students, to reinforce strengths, and to emphasize need for improvement in specific areas. The feedback concerning success of activities, other experiences, and resource speakers has been important in program planning and modification not only from one year to the next in the undergraduate program, but also for the identification of particular areas, topics or skills which need additional emphasis in subsequent quarters of the program.

The faculty has also used feedback from both formal and informal activities, not only to make program modifications, but also to resist making some changes that were considered to perhaps have a detrimental effect on the program (such as decreasing the amount of field-based work directly under the faculty's supervision, or having the students enroll in the conventional course in educational psychology, for example).

Feedback has been used to monitor some possible "plus" and "minus" effects of the program. For example, up to 1975-76, the faculty appeared to have lost some practical knowledge in the area of biological science laboratory techniques. This is probably due to the manner in which our general and special methods components are presently structured.

Feedback from graduates and from cooperating teachers with whom the faculty works over extended periods of time tends to support the concept of selective admission to our undergraduate program and to support the emphasis on continued student self-evaluation and what one faculty member terms "selective retention" in the program.

Feedback from our graduates and from their administrators has also been used when some of the faculty present introductory, recruitment-type talks to freshman and sophomore students enrolled in Ohio State's University College.

Some Key Findings and Their Implications

One key finding is that the ratings of teachers by supervisors and administrators have increased (become more positive) since 1967-68. Administrators are more positive about OSU's graduates. In addition, the graduates tend to use the teaching style which the faculty of science and mathematics education stresses: use of inquiry-oriented teaching of activities in science.

When some of the graduates have been rated less highly than others, the majority of the ratings for these individuals were expected. Faculty supervising them during their preservice program have indicated that their probable success in teaching would be limited. Data in their files provide evidence of concern for their success and potential as teachers. This finding supports a need for an emphasis on quality, both in program experiences and in students involved in the program. (This stress on quality is questioned by some faculty whose attitude is "They may go to some other school, graduate, and teach anyway, so let's take them in and try to make some kind of impression on them.")

In general, teacher attitudes regarding style used in teaching and the actual style used, regress (or drop) after the first year or so of teaching. Data through 1976 indicate that our graduates, as a group, from 1971-1974 did not regress. The program through that time appeared to have an effect that lasted on the job.

Some able graduates are having first-year-job problems. This has been a persistent problem that usually appears in those with less confidence in their abilities than others. However, first-year-job problems do not seem to appear when Ohio State faculty have been in the school working with the staff or where there has been close relationships with administrators. The implication appears to be that faculty need to work with first-year teachers and with their administrators on the job.

Present Problems

A very large problem is the need for funding for more follow-up work. This involves money for travel, for communications (mail, phone), for personnel, for computer services, for duplication of materials, and other items.

An additional problem relates to the need to wait and see what direction the College of Education follow-up efforts will take. Once this is known, then the science and mathematics faculty can decide how they can use that activity and how the College activity will affect what they are trying to do.

Considerable work has been done by graduate students either as dissertation research or for research experience. Interests of graduate students do not always follow the line of teacher education follow-up research. Such doctoral studies do cost more than many dissertations and some support for graduate students is needed.

Research Plans for the Next Year

Research plans for the coming year depends for the most part on the College of Education at OSU. It has plans to do some follow-up of this year's graduates and possibly of last year's as well. A sample of graduates will be used. The faculty of Science and Mathematics Education may be able to take a larger number of its graduates and compare to the College's sample for a validity check. At the same time the faculty, no doubt, would also collect some other data.

References

Brewington, W. C. *A study of first year secondary school science teachers who completed preservice programs at The Ohio State University*. Unpublished doctoral dissertation, The Ohio State University, 1971.

Brown, W. R. *Teacher competencies and characteristics in a science education preservice teacher education project*. Unpublished doctoral dissertation, The Ohio State University, 1972.

Cignetti, J. A. *A comparative study of the perceptions of beginning secondary science teachers in relationship to their science classroom activities, cultural attitudes and knowledge of culturally deprived students*. Unpublished doctoral dissertation, The Ohio State University, 1971.

Deamer, T. L. *The perceptions of individuals involved in a science teacher education field experience program*. Unpublished doctoral dissertation, The Ohio State University, 1973.

Erb, C. A. *A formative evaluation of an experimental teacher education project for juniors in mathematics education at The Ohio State University*. Unpublished doctoral dissertation, The Ohio State University, 1971.

Graening, J. J. *An evaluation of a secondary mathematics teacher education program emphasizing school experiences in contrasting cultural settings*. Unpublished doctoral dissertation, The Ohio State University, 1971.

Lucy, E. C. *An evaluation of a laboratory science program in a professional education course for prospective secondary science teachers at The Ohio State University*. Unpublished doctoral dissertation, The Ohio State University, 1972.

Sagness, R. L. *A study of selected outcomes of a science pre-service teacher education project emphasizing early involvement in schools of contrasting environmental settings*. Unpublished doctoral dissertation, The Ohio State University, 1970.

Swami, P. *A follow-up study for evaluation of the preservice secondary science teacher education program at The Ohio State University*. Unpublished doctoral dissertation, The Ohio State University, 1975.

TEACHER EDUCATION PROGRAM STUDY AT TENNESSEE TECHNOLOGICAL UNIVERSITY

Jerry B. Ayers
Tennessee Technological University

TEACHER EDUCATION PROGRAM STUDY AT TENNESSEE TECHNOLOGICAL UNIVERSITY

Jerry B. Ayers
Tennessee Technological University

History of the Study

Systematic evaluation of graduates of the teacher education programs at Tennessee Technological University dates back to 1964, when a mail follow-up study of the graduates was conducted for the period 1958-1963. Prior to that time, the institution's evaluation efforts had been limited to information gathered about the performance of graduates through informal contacts with the college faculty and through contacts with principals, supervisors, and superintendents. From 1964 to 1969, the College of Education conducted yearly follow-up surveys of the graduates of the teacher education programs through mail questionnaires.

The need for institutions of higher education to evaluate their teacher education programs was given added emphasis when the 1970 Standards for Accredition of Teacher Education were adopted by the National Council for Accreditation of Teacher Education. Standard 5.1 was prefaced by the following:

> The ultimate criterion for judging a teacher education program is whether it produces competent graduates who enter the profession and perform effectively. An institution committed to the preparation of teachers engages in systematic efforts to evaluate the quality of its graduates...when they complete their programs of study, and after they enter the teaching profession (NCATE, 1970, p. 12).

Standard 5.1 stated that "The institution conducts a well-defined plan for evaluating the teachers it prepares," (NCATE, 1970, p. 12). The NCATE Standards that become effective January 1, 1979, continue to give emphasis to evaluation in Standard 6.1. This standard states:

> The institution keeps abreast of emerging evaluation techniques and engages in systematic efforts to evaluate the quality of its graduates upon completion of their programs of study and after they enter the teaching profession. This evaluation includes evidence of their performance in relation to program objectives (NCATE, 1977, p. 11).

In 1970, in order to meet the NCATE standards and to gather curriculum evaluation data and sufficient information for curriculum development and reform, the College of Education at Tennessee Technological University initiated a series of systematic studies of the graduates of teacher education programs. These studies can be grouped into three broad areas: (1) systematic surveys of the graduates of the teacher education programs and employers of the graduates'; (2) special studies designed to provide specific information about a particular area of concern, such as the program for the preparation of superintendents; and (3) implementation of the Tennessee Technological University Teacher Evaluation Model. Studies under groups 1 and 2 were begun in 1970, and use of the Evaluation Model was begun in 1973. During 1977-78 the study entered its fifth year. The remainder of this paper will focus on a description of the implementation and results of the Tennessee Technological University Teacher Evaluation Model.

Overview of the Study

In a monograph prepared in 1970, for the American Association of Colleges for Teacher Education, Sandefur suggested a model for evaluating the product of teacher education programs. This model was based on generalizations made from the research literature on effective teaching and suggested procedures and instrumentation for conducting a product evaluation of teacher education programs. In 1973, the University developed and implemented a longitudinal study based on Sandefur's general model for systematic data gathering and for making decisions about programs in teacher education.

The general objectives of the program at Tennessee Technological University include the following:

1. To evaluate the objectives of the teacher education program of Tennessee Technological University through a systematic study of graduates of the program.
2. To provide information for faculty and administration concerned with teacher education programs at Tennessee Technological University in making decisions pertinent to curriculum evaluation and development.
3. To aid in the process of making long-range plans for

improving the total program of the University with particular emphasis on teacher education.

The specific objectives to be accomplished each year as a part of the study have been as follows:

1. To continue studying in a longitudinal manner those subjects who had previously participated in the application of the model.
2. To provide a descriptive profile of a sample of the previous years' graduates of the teacher education program of the University.
3. To determine relationships among selected variables as a measure of the total study.
4. To provide comparisons between the graduates of the teacher education programs of the University with those who might be considered effective teachers, as defined in the original literature on teacher education.
5. To provide effective dissemination of research data to the faculty and the administration of the University associated with the teacher education programs.
6. To provide information and suggestions for curriculum evaluation and development based on empirical research.
7. To continue to evaluate the procedures employed in the study and to make long-range plans for modifications and refinements of the basic studies.

In the fall of each year, all graduates for the past twelve months have been sent a questionnaire requesting routine demographic information about current employment, graduate's address, etc. Also, each graduate has been asked to rate his or her attainment of the major objectives of the teacher education program and to evaluate certain key courses that he or she completed. Questionnaires were sent to all individuals who received the B.S. and M.A. degrees. In turn, all individuals who were teaching in Tennessee within a 100-mile radius of the University were asked to participate in the longitudinal follow-up study (about 70% of all graduates who were teaching were within the defined limits of the study). Each year from 40-55 individuals have volunteered to participate in the study. In turn, these subjects have been followed throughout the course of the project. During the current year (1977-78), the first group of subjects will have been in the study five years. After this year it is anticipated that no further data will be collected on this first group. During 1977-78, a total of 129 graduates were participating in the study.

Data for the study have been collected each year by means of mail surveys, interviews, and observation in the graduates' classrooms. Each subject has been visited at least once each year by a specially-trained graduate assistant. The graduate assistant has observed in the subject's classroom for approximately a half day and has recorded six to eight 20-minute periods of observation using a ten-category system of interaction analysis. At the completion of all observation, the Classroom Observation Record and the Tuckman Teacher Feedback Form were completed.

The appropriate version of the Student Evaluation of Teaching (SET) was administered during each visit. The instrument has been completed by at least one class of students of each subject each year. While the students complete the SET, subjects who were participating in the project for the first year completed the California F-scale.

During the course of the day, the observer interviewed each subject with regard to opinions and ideas about the teacher preparation program of the University. Also, the observers have asked each principal to complete the Principal's Questionnaire (an instrument which allows the principal to rate the level of attainment of the objectives of the teacher preparation program by the graduate) and the Teacher Evaluation by Supervisor Form.

Pertinent data such as quality point average, National Teachers Examination scores, and other standardized measures have been gathered from permanent records. Figure 1 shows a summary of the sources of data, instrumentation, and use of data.

Data analyses have been confined largely to descriptive statistics and to the application of the t-test and analysis of variance to determine differences across years and correlational analyses to determine the relationships among the variables. Factor analysis techniques have been employed on a limited basis with certain data sets. It should be kept in mind that the data have been collected each year on as many as 150 variables.

Summary of Sources of Data, Instrumentation and Use of Data

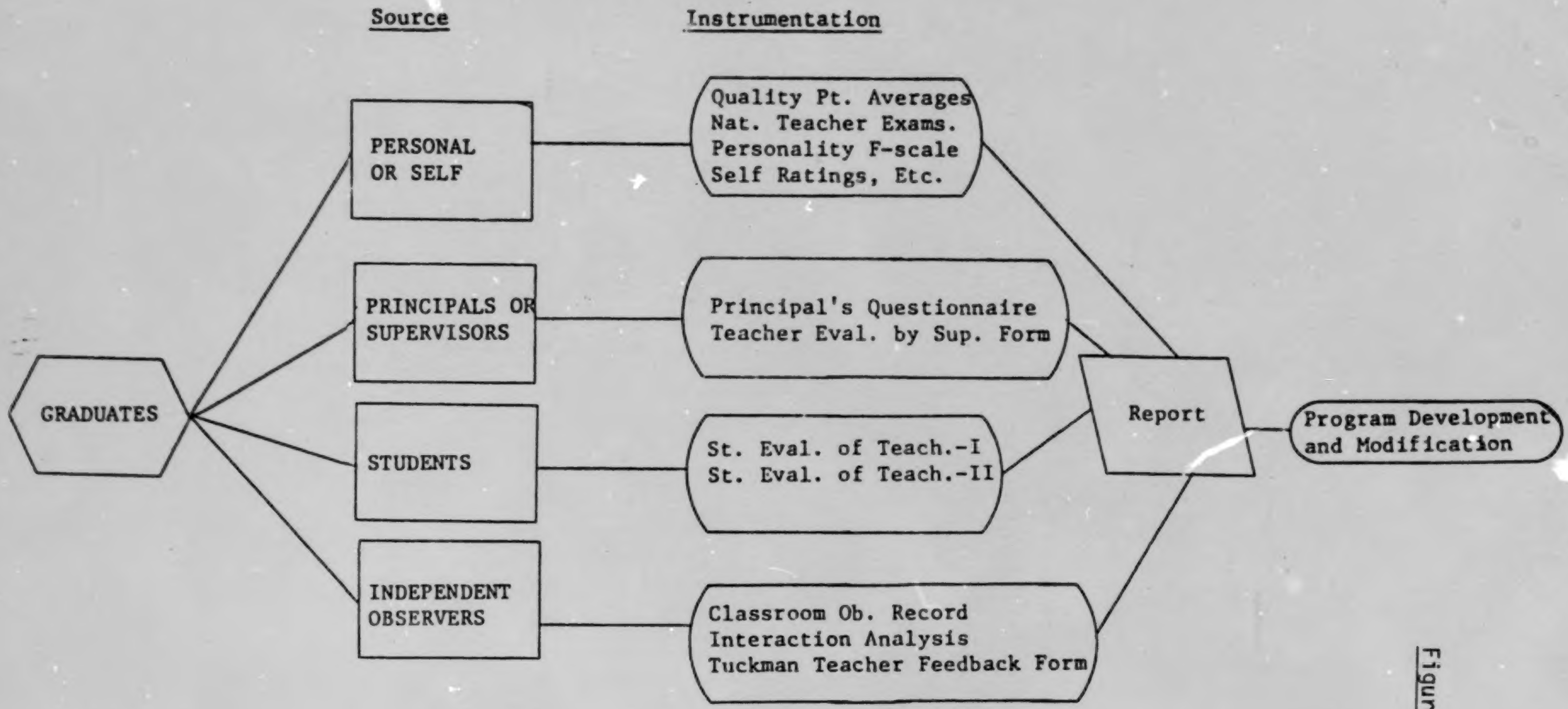

Source
Instrumentation
GRADUATES
PERSONAL OR SELF
PRINCIPALS OR SUPERVISORS
STUDENTS
INDEPENDENT OBSERVERS
Quality Pt. Averages
Nat. Teacher Exams.
Personality F-scale
Self Ratings, Etc.
Principal's Questionnaire
Teacher Eval. by Sup. Form
St. Eval. of Teach.-I
St. Eval. of Teach.-II
Classroom Ob. Record
Interaction Analysis
Tuckman Teacher Feedback Form
Report
Program Development and Modification

Figure 1

Instrumentation

Figure 1 summarizes the major instruments that have been employed in the study. A detailed description of each instrument that has been used and the manner in which data have been collected may be obtained by writing the author at the address given in the "List of Contributing Authors."

Data Base

The data base for the project is currently stored in three ways. All original data sheets and information for each subject currently in the study and for those who have dropped out have been retained in the project office. In addition, all pertinent data are stored on IBM cards, disks, and tapes in the University computer system. A computer card is available with appropriate identification number and data from each instrument for each year of the study for each subject.

Feedback

Each year, participants in the study receive a summary of their data which was gathered and also summary data for the total group. In this way, each individual can make comparisons with the total group. Participants are also invited to seek interpretation of the findings of the study from the project director. Each year 5-10 participants have requested additional information.

A report summarizing the findings of the study is prepared each year. Each report is disseminated to all individuals within the University involved in the teacher preparation programs. In addition, these reports have been disseminated widely, on request from other universities and organizations.

The results of these dissemination efforts have led to changes in the teacher preparation program of the University. Dissemination of the results outside the University has led to numerous inquiries about the implementation of the program at other institutions.

Key Findings and Implications

The development and implementation of a longitudinal study of the graduates of a teacher education program can be successfully implemented with a total investment of about $15,000 per year and has yielded results that have served to improve the teacher preparation programs of Tennessee Technological University.

Specific results of the study have indicated that the graduates who have been followed for as much as four years have changed little during the period. Few significant differences have emerged from the data. Therefore, it could possibly be concluded that a teacher's style may be developed early in the classroom career. Also, there have been few differences noted when comparing each group of first year graduates across the four years of the study. When comparisons of the data were made with that reported in the literature, the graduates appeared to be at a level comparable to or above the 50th percentile and were exhibiting many of the characteristics of good teachers. Two of the best predictors of success in teaching were level of achievement on the Professional Education Test of the NTE and overall grade point average in professional education and psychology courses. Less authoritarian graduates appeared to be overall better classroom teachers. The graduates appeared to have deficiencies in such areas as classroom control, cognitive knowledge of the science, insight into the characteristics of behavior, and effective use of community resources. Limited factor analysis of the data gathered as a part of the use of the Classroom Observation Record indicated a factor pattern similar to that obtained by Ryans (1960).

Throughout the course of the project numerous papers and monographs have been prepared. At present there is a listing of thirty-eight documents which constitute the major report efforts related to the total evaluation project. Four of these reports contain summaries of the procedures and results of the first four years of the longitudinal study of the graduates of Tennessee Technological University. Copies of all of the documents are available from the author of this paper and/or ERIC.

Problems Being Encountered

Several problems have been enountered in conducting the longitudinal study. It is anticipated that these problems can be resolved.

1. The system for computer-based management of the data is inadequate. The original system that was developed will not allow for making certain desirable analyses without complete recoding of the data.
2. The dropout rate from the project has been high. Each year about one-third of the previous year's sample of graduates have dropped out of the study for a variety of reasons.
3. It is felt that the faculty of the teacher preparation programs have not made full use of the data collected. In turn, changes and improvements in the programs for the preparation of teachers have been slow in emerging.
4. There is a need for the development of a more adequate system for evaluating the individuals who received advanced degrees (M.A. and Ed.S.), with emphasis in teaching.
5. The samples of teaching that are observed may not be representative of what actually takes place in the classroom.

Future Plans

At the present time, it is anticipated that the longitudinal study of the graduates of the teacher preparation program of Tennessee Technological University will continue on an indefinite basis. During the coming year an attempt will be made to assess the total impact of the project on the College of Education. Also, efforts will be made to redesign the system for computer storage and retrieval of data. As was noted above, this is becoming a major problem. A continual plan for monitoring the literature will be initiated, and it is anticipated that different instrumentation may be introduced into the study in the near future. During the coming year, it is planned that additional time and effort will be devoted to a more complete analysis of the existing data base, with the anticipation that answers to additional research questions will be found.

References

Ryans, D. G. *Characteristics of teachers*. Washington: American Council on Education, 1960.

Sandefur, J. T. *An illustrated model for the evaluation of teacher education graduates*. Washington: American Association of Colleges for Teacher Education, 1970.

Standards for accreditation of teacher education. Washington: The National Council for Accreditation of Teacher Education, 1970.

Standards for accreditation of teacher education. Washington: The National Council for Accreditation of Teacher Education, 1977.

EVALUATIVE STUDIES OF GRADUATES FROM TEACHER PREPARATION PROGRAMS AT OREGON COLLEGE OF EDUCATION

H.D. Schalock
The Teaching Research Division
Oregon State System of Higher Education

J.H. Garrison
G.R. Girod
K.H. Myers
Oregon College of Education

EVALUATIVE STUDIES OF GRADUATES FROM TEACHER PREPARATION PROGRAMS AT OREGON COLLEGE OF EDUCATION

H. D. Schalock
The Teaching Research Division
Oregon State System of Higher Education

J. H. Garrison
G. R. Girod
K. H. Myers
Oregon College of Education

Like most institutions that prepare teachers, Oregon College of Education (OCE) has had an uneven history of evaluation of its graduates. During the early 1960's, and as recently as 1967, the College regularly interviewed graduates at their employment sites and reported summaries of these interviews to appropriate faculty groups (Kersh, et al., 1976). The interview procedure was informal, however, and faculty members consistently reported that the information contained in these reports had already been communicated to them by other means. Since a quarter of a faculty person's time was allocated each year to carrying out these interviews, and considerable expense was involved in off-campus travel, the practice was discontinued in 1967 because the benefits did not appear to justify the cost. Prior to 1960, the College periodically carried out questionnaire surveys of its graduates who were teaching, and between 1972 and 1974 the College, through its registrar, conducted an extensive survey of the 1967 and 1970 graduating classes who received baccalaureate degrees in education (Kenyon, 1976).

In 1972, with the initiation of a competency-based and field-centered approach to the preparation of its elementary teachers on an experimental basis, Oregon College of Education (OCE) had need for better information regarding the effectiveness of the graduates of its various teacher preparation programs. The experimental program was somewhat more costly to operate than the previous program (Schalock, Kersh and Garrison, 1976), and evidence had to be obtained as to benefits received. Informal evaluations at the conclusion of the 1972-73 experimental year led to the decision on the part of the College to adopt the program as its regular approach to the preparation of elementary teachers, but this decision was accompanied by a parallel decision to engage in systematic follow-up studies as a basis for program evaluation and adaptation.

Another purpose for initiating a systematic program of follow-up at OCE was the information it would provide for use in research, underway at the College and in the Teaching Research Division, on the ability to predict the effectiveness of teachers. Data obtained through evaluative follow-up studies of first-year teachers, and more experienced teachers, would provide the "dependent" measures needed in the research program to determine whether early predictors of success in teaching were effective. Early predictors made available through the comptency-based and field-centered elementary program included performance in lesson teaching, performance in two to five days of full responsibility teaching, and performance in student teaching.

The purpose of this paper is to describe the various follow-up studies that have been done at OCE since 1974, and the projected plans for follow-up in the years ahead. The report of findings from these studies will be kept brief, since they are primarily of interest to OCE. Attention will be given, however, to "lessons learned" from each study about doing follow-up studies, and to the costs and benefits that have been found to accompany such studies. Attention will also be given to the use made by other institutions in Oregon of the follow-up methodology developed at OCE and the Division of Teaching Research, and of steps now underway in Oregon to develop a "common core" methodology that can be used by all institutions in the state in carrying out evaluative studies of the graduates of their teacher preparation programs.

The methodology that has been used in the various studies and reports of findings are available upon request. (See "List of Contributing Authors" at the end of this volume.)

The November, 1974 Planning Conference

In November, 1974, a conference was held on the OCE campus to identify the elements and procedures thought at that point to be critical to successful follow-up studies of graduates of teacher preparation programs.

The conference involved persons from colleges throughout the nation who had implemented such studies, as well as students and staff from OCE and personnel from local schools. The conference led to a decision by OCE faculty and administrators to try out three alternative approaches to the collection of follow-up information on graduates. The alternatives were:

1. A mail-only design;
2. A telephone-mail design; and
3. An on-site visitation design, preceded by both telephone and mail contact.

The mail-only and telephone-mail designs differed only in the approach used to obtain the kind of information desired. The on-site visitation design provided the added advantages of observation and interview data, and the opportunity to clarify information prepared by the graduate and the graduate's principal prior to the site visit.

The 1975 Pilot Study

This study was designed primarily as a test of the alternative methodologies outlined above. It was also designed to provide preliminary information about the effectiveness of graduates from the new elementary program, but this was secondary to the purpose of methodological development and testing.

Methodology

The observation procedures employed in the study by the OCE observer represented an extension of the procedures used to assess the performance of prospective teachers in the elementary preparation program. Questionnaire and interview procedures focused on the contribution of the elementary program to present level of competence, the attitudes the graduate held toward teaching and being a teacher, perceived strengths and weaknesses in the preparation program, recommendations for strengthening the program, etc. Both the graduate and the graduate's principal were interviewed, and both provided judgments as to the competence of the graduate in performing the functions expected of them in their present positions. These judgements were provided on instruments that reflected the same teaching functions assessed by the OCE observer.

The survey instruments developed for the mail-only and the telephone-mail designs were derived from and corresponded as closely as possible to the instruments used in the on-site design.

Sample

A ten percent sample of 1974 graduates from the elementary program that were employed as first year teachers were selected for follow-up through each design. The sample consisted of 45 teachers, 15 in each of the three design groups. The sample was selected on the basis of convenience of access rather than the principles of random sampling.

Findings

The information gained from the three competing designs was evaluated with respect to three issues: (a) the kind of participation obtained from graduates and their principals; (b) the kind and quality of information obtained; and (c) the relative benefits of the various designs in comparison to their costs. Major findings included:

1. All teachers and principals contacted for participation in the study that involved the site visit participated; whereas less than one-half of those contacted through either of the competing designs participated.
2. The information yielded, both in quantity and quality, was much higher for the site visit design than for the competing designs.
3. The "public relations" value gained from the on-site visitation, and the value of these visitations to college faculty, were much greater than anticipated.
4. The cost of the site visit design was much greater than either of the other two designs.

On the basis of these findings the on-site methodology appeared to have clear advantage over the competing methodologies, even though it was considerably more expensive. Because the sample of graduates contacted in the pilot study was small and non-random, however, and because returns were poor from the mail-telephone and mail-only methodologies, it was not possible to reach firm conclusions about the costs and benefits associated with the three methodologies. As a consequence the decision was made to test the two most promising methodologies another year. These were the on-site and the telephone-mail methodologies, as refined on the basis of their use in the 1974-75 pilot study.

The substantive results found in the pilot study have been reported in detail by Garrison and Girod (1975).

Lessons Learned

1. It is possible within a single day at a school to (a) observe a graduate teach on at least two separate occasions for a full class period; (b) carry out interviews with the graduate, the graduate's principal, and pupils and colleagues of the graduate with sufficient care that a reasonably detailed and "cross-validated" check on the competence of the graduate as a first year teacher can be gained; (c) obtain the graduate's and the principal's perception of the strengths and weaknesses of the preparatory program.

2. The instrumentation and scheduling needed to accomplish these various tasks require careful development.

3. Procedures used in contacting the graduate and the graduate's principal about taking part in the program are as critical to the success of a follow-up study as the instrumentation and scheduling procedures used in data collection.

4. Both graduates and their principals are willing to and able to prepare materials in advance of the college observer's visit for use in the visitation or for analyses subsequent to the visitation.

The 1976 Study of First Year Teachers at the Elementary Level

This study of first year teachers who were graduates of the Elementary program at OCE built upon the methodologies and procedures tested in the 1974-75 pilot study. The study served two purposes. First, it provided a more carefully controlled test of the costs and benefits associated with the on-site and the telephone-mail methodologies. Second, it provided a careful and thorough study of the first year teaching effectiveness of graduates of the competency-based and field-centered program at OCE.

Methodology

Seven major sets of data were to be collected through both the on-site and telephone-mail methodologies:

1. A judgment by the principal about the interpersonal relationships, involvement and leadership ability of the OCE graduate in relation to other teachers in the school.
2. A judgment by the graduate, the graduate's principal, and the OCE observer in the on-site study as to the graduate's competence in the performance of particular teaching functions, including the achievement of a variety of content and attitudinal outcomes in children.
3. A description by the graduate of the major pupil outcomes pursued over a two or three month period of time, and evidence of their achievement.
4. A judgment by the graduate, the graduate's principal and the OCE observer as to the difficulty of the graduate's school setting as a context in which to teach.
5. A judgment by the graduate and the graduate's principal as to the graduate's attitudes toward teaching and being a teacher.
6. A judgment by the graduate as to the contribution of the OCE elementary program to the level of competence he/she currently possesses as a first year teacher.
7. The perception of both the graduate and the graduate's principal of weaknesses in the current elementary program at OCE, and recommendations for improvement.

Two additional sets of data were to be collected through the on-site methodology by the OCE observer:

1. A judgment from the graduate as to how typical the teaching setting was during the OCE faculty member's observation.
2. Interpretive comments about the graduate's performance as a teacher on the basis of casual observations or discussions with pupils and/or colleagues during the course of the on-site visitation.

110

Sample

A stratified random sample of 24 graduates of the elementary program serving as first year teachers in schools that fell within a radius of fifty miles of the OCE campus was contacted for study through the on-site methodology. The sample was stratified on the basis of teaching assignment (primary or intermediate) and school setting (rural or urban). The sample sought and obtained in the study is illustrated in Table 1. The final sample is shown in parentheses.

Table 1

Sample Design for the 1976 On-Site Study of First Year Elementary Teachers

	Primary	Intermediate
Rural	6 (6)	6 (6)
Urban	6 (5)	6 (5)

Of the remaining pool of graduates from the OCE elementary program serving as first year teachers, thirty were selected for follow-up through the mail-telephone methodology. Twenty-four of the thirty were selected on a stratified-random sample basis identical to that followed in the on-site methodology, though they were not limited to the fifty mile radius imposed on the sample drawn for the on-site study. All, however, were teaching within Oregon. Six additional graduates who were teaching outside Oregon were included in the mail-telephone sample. This portion of the sample was not stratified in terms of teaching assignment or school setting. The sample of graduates sought and obtained in the mail-telephone study is shown in Table 2. The final sample appears in parentheses.

Table 2

Sample Design for the 1976 Mail-Telephone Study of First Year Elementary Teachers

	In-State		Out-of-State
	Primary	Intermediate	
Rural	6 (4)	6 (5)	6 (1) non-stratified
Urban	6 (2)	6 (2)	

Findings

Essentially the same results were obtained in the 1976 study as in the pilot study regarding the relative yield from the two methodologies tested. The on-site design provided much richer data, and data that could be viewed with greater confidence, than data coming from the mail-telephone design. The results obtained from the two studies simply in terms of data returns summarized in Table 3.

The costs of carrying out the two studies were markedly different. The costs of the on-site study was approximately $10,000; whereas the cost of the telephone-mail study was approximately $2,000. A cost break-down for the two designs appears in Table 4. Personnel is divided into two categories: OCE Personnel and TR (Teaching Research Division) Personnel.

The substantive findings from the study were most encouraging, in that 19 of the 22 graduates studied were judged by all concerned to be functioning as fully competent first year teachers. The three weakest graduates studied were judged to be less than fully competent with respect to some functions they were expected to perform as teachers, for example, assessing learning outcomes desired in students; but none were in danger of losing their jobs. Furthermore, eight of the twenty-two graduates studied were judged by their principals to have assumed positions of leadership within their schools, and all twenty-two of the graduates were planning to continue their teaching careers. Additional findings of general interest include:

1. Graduates who had classrooms that were judged to be complex/difficult contexts in which to teach were judged consistently as less competent in the performance of teaching functions than graduates whose classrooms were judged to be easy contexts.

2. Graduates identified as holding high leadership/status positions within their schools by their principals reflected consistently higher levels of competence in the performance of teaching functions than graduates judged by their principals as having relatively low leadership status positions; but all graduates who reflected high levels of competence had not assumed positions of leadership.

3. There was essential agreement between the OCE observer, the graduate and the graduate's principal as to which of the

Table 3

A Summary of the Primary Data Sets Collected Through the Two Methodologies

DATA SET	ON-SITE	MAIL-TELEPHONE
A judgment by the principal of the interpersonal relationships, involvement and leadership ability of the OCE graduate.	22	12
Judgments as to the competence in the performance of particular teaching functions, including the achievement of a variety of content and attitudinal outcomes in children:		
by the OCE Observer	22	NA
by the Graduate's Principal	22	12
by the OCE Graduate	22	12
A description of pupil outcomes pursued over a two or three month period of time, and evidence of their achievement.	5	2
Judgments of the difficulty of the graduate's classroom as a context in which to teach:		
by the OCE Observer	12	NA
by the Graduate's Principal	15	12
by the OCE Graduate	22	12
Judgments about the attitudes of the graduate toward teaching and being a teacher:		
by the Graduate's Principal	22	11
by the OCE Graduate	22	12
A judgment as to the contribution of the OCE elementary program to the level of competence currently possessed by the graduate as a first year teacher.	22	12
Perceptions of weaknesses in the OCE elementary teacher preparation program, and recommendations for their improvement:		
by the OCE Graduate	22	11
by the OCE Graduate's Principal	22	11

Table 4

Costs Associated with the 1976 Follow-up Study of Graduates of the Elementary Teacher Preparation Program at OCE

Category of Costs	Telephone-Mail	On-Site
OCE Personnel		
--Design, Instrumentation, Sample Selection, Initial Contacts	TAPORA*	TAPORA
--Data Collection	$ 500	$4,500
--Data Reduction, Analysis, Interpretation, Reporting	100	500
TR Personnel		
--Design, Instrumentation, Sample Selection, Initial Contacts	200	1,200
--Data Collection	NA	NA
--Data Reduction, Analysis, Interpretation, Reporting	960	3,000
Materials and Supplies	38	75
Duplication	124	245
Travel and Per Diem	NA	275
Communications	75	50
Totals**	$1,997	$9,845

* TAPORA: Treated as part of regular assignment.

** Overhead (49.4% of salaries), personnel benefits (15% of salaries) and costs for reproduction of final reports are not included in this figure.

graduates were not fully competent and which were unusually competent.

The substantive findings from this study have been reported in detail by Schalock, Girod and Garrison (1976).

Lessons Learned

1. The data received from mail-only, or a combination of telephone and mail designs, will probably always be of questionable validity.
2. The on-site methodology developed and tested in the 1974-75 pilot study and again in the 1976 study, is a productive and economically feasible approach to the collection of evaluative information on graduates of teacher preparation programs, though it is approximately five times as expensive as mail surveys or mail-telephone designs.
3. Graduates were unwilling or unable to provide information about pupil outcomes pursued over a two or three month period of time, and evidence of the achievement of these outcomes.
4. The analysis and reporting of data obtained through a methodology that is as productive of information as the on-site methodology used at OCE becomes an enormous task, and resources must be set aside for it if the results of a follow-up study are to be of value. Ways also must be found to tailor the reports coming from such studies to the needs of the various audiences served by the reports, and to keep the reports as short and simple as possible.

The 1976 Study of First Year Teachers at the Secondary Level

After studying the results of the 1975 pilot study, and in light of the 1976 study that was to be conducted by the elementary faculty, the faculty of the secondary preparation programs at OCE decided to apply the on-site methodology in a follow-up study of their recent graduates who held first year teaching positions. The decision was made to use the more complex and costly methodology, because the greater richness of data produced and the greater confidence that could be placed in the data seemed to warrant its cost. Adaptations were made in the methodology to fit the demands of the secondary program, but in all important respects it was the same methodology as employed in the 1976 study by the elementary division.

Sample

The stratified-random sample consisted of thirty-four first year teachers who were graduates of the secondary program and who were teaching within a fifty mile radius of the OCE campus. The sample was stratified on the basis of subject matter specialization.

Using the number of students who graduated from each subject matter area during 1974-75 as the basis for calculating sample size, a 17% sample of all graduates was drawn for purposes of the study (34 from a total of 191 graduates). The sample represents approximately 30% of the total number of graduates placed in teaching positions for the 1975-76 school year. With the exception of Biology and Health, four graduates were drawn from each subject matter area in which secondary teachers are prepared. The sample initially drawn for the study and the final sample obtained are shown in Table 5. The final sample obtained appears in parentheses. The percentages shown represent the percent of graduates observed to the number of students graduated in 1975, from each of the subject matter specializations.

Table 5

Sample Design for the 1976 Follow-up Study of First Year Secondary Teachers Who Were Graduates of the OCE Secondary Programs

Art	Biology	Health & Phys. Ed.	Humanities	Math	Music	Physical Science	Social Studies
4	3	7	4	4	4	4	4
(2)	(3)	(6)	(4)	(3)	(4)	(3)	(4)
9%	33%	14%	13%	38%	19%	27%	11%

Findings

As anticipated, the on-site methodology developed and tested within the elementary program worked, with adaptations, at the secondary level. The graduates and their principals were able to apply the methodology, as were faculty from the Department of Education and faculty from academic departments who held joint appointments in education. Moreover, the cost of implementing the follow-up study at the secondary level was less than at the elementary level, since college supervisors assumed responsibility for the on-site visitations as part of their regular job assignment. An estimate of costs associated with the study placed "out-of-pocket" expenses at approximately $3,500.

Substantively, the results of the secondary study were as satisfying to the College as were the results of the Elementary study, though five of twenty-nine graduates studied were judged to be less than fully competent as teachers. Competence in the performance of teaching functions did not appear to vary appreciably by high school or junior high teaching assignment, or by the rural-urban setting of a school. Competence in the performance of teaching functions did appear to vary, however, by content area. Graduates teaching in several of the content areas were judged to be noticeably less competent as a group than graduates from other content areas. Another finding of interest was the fact that graduates of the secondary programs had not emerged as leaders within their respective schools to the same extent as had graduates of the elementary program, but this was felt to be as much a matter of contextual differences as it was the quality of graduates or the nature of the preparatory program.

Lessons Learned

1. It is hard to draw inferences about the effectiveness of preparation programs from a sample of two or three graduates of that program.

2. It is even harder to draw inferences about preparation programs when graduates studied are transfers to the program at the end of their sophomore or junior years, or come to the program after completing their undergraduate degree elsewhere--a discovery which was made only after the analysis of the data had been completed and implications were being considered.

3. There has to be an adaptation of the content and emphasis in a follow-up methodology to accomodate differences in content and emphasis within the various subject matter areas involved.

4. Conducting follow-up studies on graduates of secondary preparation programs is considerably more complex than conducting follow-up

studies on graduates of elementary programs.

The 1977 Study of First Year Teachers at the Secondary Level

Because of the sample limitations in the study just reviewed, a decision was made to replicate the study in 1977. The methodology remained essentially unchanged, though a number of the instruments were refined on the basis of further use in the program or as a result of use in the 1976 study. Both the competency assessment instruments used by the OCE observer and the instruments describing the setting as a context within which to teach, saw major revision.

Sample

In keeping with the initial study, a stratified-random sample of thirty-four first year teachers who were graduates of OCE secondary programs were asked to take part, with stratification occurring on the basis of subject matter specialization. In contrast to the 1976 study, where only graduates teaching within a 50-mile radius of the College were included, the 1977 study included graduates teaching within a 100-mile radius of the campus.

Using the number of 1976 graduates from each subject matter area as a basis for calculating sample size, a twenty-nine percent sample of all graduates placed in teaching positions during the 1976-77 school year were drawn (34 from a total of 117 employed). The sample drawn initially for the study, and the final sample obtained (shown in parentheses) are illustrated in Table 6. The percentages shown represent the percent of graduates observed to the number of students graduated in 1976 from each of the subject matter specializations.

Table 6

Sample Design for the 1977 Follow-up Study of First Year Teachers Who Were Graduates of the OCE Secondary Program

Art	Biology	Health & Phys. Ed.	Humanities	Math	Music	Physical Science	Social Studies
4	2	6	5	3	3	5	6
(3)	(2)	(5)	(5)	(3)	(3)	(5)	(6)
38%	67%	23%	28%	60%	38%	100%	32%

Findings

Substantively, the results of the study essentially paralleled the results obtained from the previous study, though differences did occur within subject matter areas. Several of the content areas, whose graduates were judged to be particularly strong teachers in the 1976 study, continued to have their graduates reflect this pattern of excellence. Several subject matter areas that had graduates judged to be relatively weak in the 1976 study, however, had graduates who were judged to be relatively strong as teachers in the 1977 study. Other subject matter areas had a reverse pattern occur. There was no consistency between the first and second study as to the content areas where the weaker graduates were found.

Findings relative to the methodology essentially substantiated earlier findings; namely, the methodology provides a feasible and acceptable set of procedures by which to obtain a great deal of information about the performance of graduates of secondary teacher preparation programs, as well as information about the perceived strengths and weaknesses of those programs and how these weaknesses might be overcome.

Lessons Learned

1. One is limited in drawing conclusions about the effectiveness of a preparatory program on the basis of evaluative studies of a small sample of graduates in a particular year.

2. Even though there is some shifting about in the relative performance of graduates of various subject matter departments, as reflected in data from two independent studies, there appears to be consistency across most departments.

3. Inferences probably can be drawn about the effectiveness of preparation programs at the secondary level from relatively small sample studies if there is consistency of evidence across at least three years of research findings.

4. It is easy to underestimate costs involved in conducting follow-up studies. Estimated costs for the 1976 study of graduates of the secondary program were placed at $3,500. Detailed cost estimates for the 1977 study placed the cost at $5,199 in out-of-pocket expenses; $12,461 in contributed time by faculty and administrators from the College; and $3,296 in contributed time by the schools.

The 1977-78 Longitudinal Study of Graduates from the Secondary Preparation Program in Humanities

As an accompaniment to the 1977 study of OCE graduates from secondary programs, a study was undertaken that involved all of the 1976 graduates in teacher education from the Department of Humanities (Nilson, 1977). Fifty students graduated from the Department in 1976 with teaching certificates. Twenty-one of these graduates obtained positions as first year teachers in 1977. Six of these in turn were evaluated in the regular follow-up study. The remaining fifteen were evaluated as part of the Nilson study. These twenty-one graduates are being evaluated as second year teachers this spring.

While the results of this study are not yet in, preliminary findings have been most informative. These include:

1. Of the twenty-nine graduates who did not obtain teaching positions in 1977, seventeen obtained such positions in 1978.
2. Five additional graduates obtained positions as substitute teachers, or teachers' aides.
3. The six graduates included in the regular follow-up study were not particularly representative of the twenty-one graduates studied in 1977 as a whole, or of the communities in which graduates were teaching, even though they were drawn randomly from the pool of 21 graduates who held first year teaching positions.
4. Graduates who were having most difficulty in their jobs tended to be in trouble largely because of cultural background or values that were at odds with the culture or values of the community in which they were teaching, not because they lacked competence as teachers.

All these findings constitute "lessons learned" about conducting follow-up studies, and need to be taken into account in designing a long-term follow-up strategy.

The 1978 Study of Third Year Teachers Who are Graduates of the Elementary Program at OCE

The Elementary Program is currently conducting a study of graduates evaluated in the 1976 study who are now third year teachers. Eighteen of the twenty-two graduates studied in 1976 are still teaching in Oregon, and all have agreed, along with their principals, to take part in the third year study. These instruments may be obtained by writing to the authors (see "List of Contributing Authors").

The Long-Term Plan for Evaluative Studies of OCE Graduates from Teacher Preparation Programs

Plans now call for the systematic evaluation of graduates from the elementary and secondary preparation programs at OCE, as first, third, fifth and ninth year teachers. The design is longitudinal, in that teachers evaluated as first year teachers will be followed as third year teachers, fifth year teachers, etc. The design also calls for the selection of teachers to be stratified on the basis of predictions as to their likely success. This will be based on their performance in the undergraduage program at OCE. A three-level stratification is called for: (a) those predicted to be unusually successful as teachers; (b) those predicted to be competent but not outstanding as teachers; and (c) those falling somewhere between these extremes. This design is shown schematically in Table 7.

Table 7. Long-Term Plan for Evaluation Studies

	Years in Teaching			
	1	3	5	9
Graduates of the Elementary Program				
--Predicted to become outstanding as teachers				
--Predicted to have a chance of becoming outstanding				
--Predicted to be competent as teachers, but not outstanding				
Graduates of the Secondary Program				
--Predicted to become outstanding as teachers				
--Predicted to have a chance of becoming outstanding				
--Predicted to be competent as teachers, but not outstanding				

Evaluative follow-up studies of graduates from the various advanced and specialist certification programs offered by the College, and from students who enter the elementary/secondary programs with undergraduate degrees from elsewhere, are also to be studied for evidence of program effects, but these studies will take different forms and foci than studies projected for graduates of the regular elementary and secondary undergraduate programs.

The adoption of a longitudinal, stratified sampling approach to follow-up studies at OCE rests on a number of assumptions and long-term aims. These include:

1. The assumption that competence as a teacher grows with experience and varies by the context in which teaching occurs, and that if there is pattern in this change it can be determined only through a longitudinal design.
2. The fact that we do not know whether we can predict excellence in teaching early in a teacher's career, and that we will never know whether we will be able to do so in the absence of a stratified, longitudinal design of the kind proposed.
3. The fact that we do not know how to identify in early stages of teaching the teachers that are likely to emerge as leaders among their colleagues in schools, and that we will never know whether we will be able to do so in the absence of a design of the kind proposed.
4. The need to be able to determine the long-term effects of planned variations in the preparation programs at OCE, both in terms of maintaining a variation once implemented and determining when existing programs need to be modified to overcome identified weaknesses.

All of the above are in keeping with the design of the teacher preparation programs at OCE as contexts for research on teacher and program effectiveness. Without longitudinal studies on the effectiveness of graduates as first year and experienced teachers, research on either program effectiveness or the ability to predict success in teaching simply is not possible.

Use Made of the Methodology by Other Oregon Institutions

Four other institutions that prepare teachers in Oregon have adopted the methodology developed at OCE and Teaching Research for use in their own follow-up studies. These are Eastern Oregon State College, Southern Oregon State College, the University of Oregon, and Oregon State University. Eastern and Southern have used the methodology in studying graduates in both their elementary and secondary programs; the University of Oregon has used it in evaluating graduates of their resident teacher program; and Oregon State University began to use it in the spring of 1978, in evaluating graduates of its elementary program. A report comparing the results of the OCE study on graduates of its elementary program to graduates of the University of Oregon's resident teacher program is presented in Kehl (1978).

The OCE and TR methodology also has been used as a point of departure in developing a "common core" of measures and procedures that can be used by all teacher preparation institutions in the state in studying the performance of their graduates. This common core of methodology is sufficiently far along in its development that there are plans for its field testing in the spring of 1979. Studies undertaken in the spring of 1978, at the University of Oregon, and the study underway at OCE on graduates of the elementary program as third year teachers, also are serving to pilot some of the elements within the proposed methodology.

Once this common core of methodology is available, it is anticipated that studies on the costs and benefits of alternative teacher preparation programs within the state will be undertaken. It is anticipated further that evidence coming from these studies will be used as a basis for initiating an empirically-based approach to the improvement of the preparation of teachers in Oregon. Research designs appropriate for such studies, and the conditions that must be present for such studies to be undertaken productively, have been outlined by Schalock in a monograph published by the Multi-State Consortium on Performance Based Teacher Education, entitled Closing The Knowledge Gap: CBTE Programs As A Focus Of and Context For Research in Education (1975).

References

Garrison, J. H., & Girod, G. R. Report of the 1975 OCE follow-up program: Elementary education graduates. Monmouth, Oregon: Oregon College of Education, 1975.

Kehl, E. The resident teacher education program: A pilot evaluation study. The University of Oregon, 1978.

Kenyon, S. J. Effectiveness in fulfilling the stated goals of a teacher education program as evaluated by graduates of that program. Ed.D. Thesis, Oregon State University, 1976.

Kersh, B. Y., et al. Institutional report for the national council for accreditation of teacher education. Monmouth, Oregon: Oregon College of Education, 1976.

Nilson, M. E. A longitudinal study of 1976 graduates from the humanities department involved in teaching. A proposal for doctoral dissertation, The University of Oregon, 1977.

Schalock, H. D., Girod, G. R., & Garrison, J. H. Summary of the 1975-76 follow-up data on first year teachers who are graduates of the OCE elementary teacher preparation program. Oregon College of Education, 1976.

Schalock, H. D., Kersh, B. Y., & Garrison, J. H. From commitment to practice: The OCE elementary teacher education program. Washington, D.C.: The American Association for Colleges of Teacher Education, 1976

Schalock, H. D. Closing the knowledge gap: CBTE programs as a focus of and context for research in education. Multi-State Consortium on Performance Based Teacher Education, Syracuse University, 1975.

Schalock, H. D., et al. Summary of the 1975-76 follow-up data on first year teachers who are graduates of the OCE secondary teacher preparation program. Oregon College of Education, 1976.

Schalock, H. D. et al. Summary of the 1976-77 follow-up data on first year teachers who are graduates of the OCE secondary teacher preparation program. Oregon College of Education, 1977.

EVALUATION OF INSTRUCTIONAL SYSTEM CHARACTERISTICS OF THE PROFESSIONAL TEACHER PREPARATION PROGRAM AT THE UNIVERSITY OF HOUSTON

Wilford A. Weber, James M. Cooper
University of Houston

EVALUATION OF INSTRUCTIONAL SYSTEM CHARACTERISTICS OF THE PROFESSIONAL TEACHER PREPARATION PROGRAM AT THE UNIVERSITY OF HOUSTON

Wilford A. Weber and James M. Cooper
University of Houston

Introduction

The Professional Teacher Preparation Program (PTPP) is the undergraduate teacher education program of the University of Houston's College of Education. The PTPP is designed for students pursuing baccalaureate degrees and provisional certification as elementary school teachers, secondary school teachers, or teachers of all levels in art, health and physical education, or music. Thus, the PTPP is a college-wide effort to develop and operate an integrated and comprehensive, competency-based instructional system for the training of undergraduate teacher education students.

The PTPP traces its beginning to 1966 when the faculty of the College of Education committed itself to innovative approaches for the design and operation of professional teacher education programs. This effort was greatly enhanced by a 1970 grant from the Trainers of Teacher Trainers (TTT) Project and by subsequent grants from the Texas Teacher Center Project and the Teacher Corps Program. By the end of the first year of the TTT Project, the design of a competency-based teacher education program had been completed. This program was field tested with sixty-four prospective teachers--half of whom were elementary education majors and half of whom were secondary education majors--during the 1971-72 and 1972-73 academic years. The program was revised on the basis of data from this field test, and an additional 121 students participated in the revised program during the 1972-73 and 1973-74 academic years.

Having assessed the impact of the program during these two field tests, in the spring of 1973, the faculty voted to utilize the competency-based approach in all undergraduate teacher education programs in the College, beginning with the fall semester of that year. Since that time, effort has been directed toward making the entire undergraduate teacher education program a competency-based instructional system.

Assessment of the program's first year of full operation suggested the need for far better intergration and coordination of the program's components. Systemic development had been lacking. A rather piecemeal effort had resulted in many good, but ill-fitting program "pieces." Therefore, in the summer of 1974, in order to enhance the program's effectiveness and efficiency, a management system was created. The Program Development and Implementation Council (PDIC) became the management team responsible for coordinating and facilitating the conceptualization, design, development, implementation, operation, and evaluation of the PTPP.

The central task of the PDIC has been and continues to be one of program coordination and facilitation. Responsibility and accountability for the design and delivery of the various instructional components of the PTPP have rested with various operating units within the College. The PDIC has undertaken extensive program development activities in an effort to conceptualize and design an integrated, comprehensive, and effective instructional system. Much of this work took place during the 1974-75 academic year. This included the specification and, in April of 1975, the formal adoption of twelve assumptions regarding those instructional system characteristics the faculty believed were essential to the operation of an effective teacher education program. Thus, the PTPP has been built on the assumption that an instructional system for the preparation of teachers is more effective when it is:

1. Competency-based. Competencies to be demonstrated by the student are made explicit, the criteria to be applied in assessing the student's competencies are made explicit, and the student is held accountable for meeting those criteria. The emphasis is on the demonstration, not the acquisition, of those competencies specified as program expectations.

2. Campus-centered and field-oriented. Students are provided with opportunities to experience instruction and to demonstrate competence both on campus and in field sites, depending upon the nature of the particular instructional or assessment activity and a determination as to where that activity can most effectively and efficiently take place.

3. Role-model based. A conceptual model for the teacher's role is used as a basis for identifying and specifying those competencies students are expected to demonstrate.

4. Criterion-referenced. Assessment procedures are designed to determine whether or not a student has demonstrated competency at or above the level of mastery specified; the competency of each student is judged on the basis of predetermined criteria and is not determined through a comparison involving other students.

5. Pluralistic. There is conscious acceptance of the notion that no one philosophy of instruction has been proved best; consequently, a variety of divergent views are posited as untested assumptions which the system must test.

6. Humanized. Program design and operation recognize the dignity and worth of each individual so that each perceives that he or she is being treated as one of worth.

7. Personalized. The student is provided with instruction which takes into account his or her uniqueness.

8. Modularized. The delivery of instruction is accomplished through the utilization of instruction modules; instructional modules are sets of learning activities--rationale, objectives, prerequisites, pre-assessment procedures, learning alternatives, post-assessment procedures, and remedial procedures--which are intended to facilitate the learner's acquisition and demonstration of a particular competency or set of competencies.

9. Multi-institutional. The design and operation of the teacher education program is a responsibility shared by colleges, public schools, and the organized teaching profession.

10. Systemized. The systems approach is used in program design and operation.

11. Regenerative. The program is an open system capable of continuous revision on the basis of constructive data supplied by sound formative and summative evaluation procedures.

12. A single system with alternatives within it. The program is a single instructional system which constitutes an integrated, comperehensive whole; however, the system does accomodate a divergence of viewpoints within its confines.

While the 1974-75 academic year was characterized by efforts of a program conceptualization and design nature, the 1975-76 academic year was marked by a move from program design to program development. These program development activities included a conscious effort to develop an instructional system which exemplified the twelve operational characteristics which had been adopted.

Full implementation was undertaken during the 1976-77 academic year. During this period, program development had largely given way to full program implementation. Operation of the program during the 1976-77 academic year, in effect, constituted a field test, as this period was the first time that the college had tried to "put together all of the pieces" of the PTPP. Program development had reached that stage when it became possible--and desirable--for evaluation processes to describe and analyze those "pieces" and the whole they constituted. Thus, program evaluation was viewed as the next program development task to be undertaken.

In anticipation of the need for program evaluation, in the spring of 1976, the PDIC appointed a task force composed of seven faculty members and chaired by the associate dean for undergraduate studies. The task force was given responsibility for designing and carrying out a comprehensive evaluation of the PTPP. The major purpose of that evaluation was, and continues to be, to provide data from which to make informed decisions regarding program refinement and improvement. This paper describes one aspect of that effort: an evaluation of the extent to which the twelve instructional system characteristics which had been adopted in 1975 were operationalized by the program during the 1976-77 academic year.

Goals

A first step in the task force's efforts to design and carry out a comprehensive evaluation of the PTPP was the specification of the goals and objectives the evaluation process was to achieve. It was intended that the evaluation process should:

1. Provide information which would be used to support the instructional system's efforts to facilitate student competency attainment.
2. Provide information concerning student progress, student achievement, and program effectiveness.
3. Provide information which would be useful in promoting closer and more productive communication with students.
4. Provide information which would be responsive to the

expressed data needs of faculty who are implementing the program.

5. Provide data which are timely, easily obtained, and credible.

6. Provide an information base for decision-making related to program revision.

Given these goals, it was determined that a task force sub-group consisting of three persons should undertake an evaluation of the extent to which the PTPP instructional system was exhibiting the twelve operational characteristics agreed to during the program's conceptualization. It seemed to be important to describe the instructional system and determine the extent to which it was, or was becoming, what it was intended to be. This seemed to be an especially important task for four major reasons. First, program designers had spent considerable effort in conceptualizing and specifying the instructional system characteristics they believed were essential to program effectiveness, and it appeared to be of benefit to determine the extent to which they were successful. Second, while much had been written and said about competency-based teacher education, it appeared that few competency-based teacher education programs existed. Indeed, there were educators who seemed to express doubts concerning the viability of creating such programs. It was assumed that there was benefit to be gained by testing that assumption. Third, if anything was to be said about the program's effectiveness with regard to its ability to facilitate student competence, it seemed quite essential that the nature of that program's instructional system be thoroughly described and understood. Fourth, if the program was to be improved, there was a crucial need for information about the nature of its instructional system.

As one examines the reasons given above, it is essential to understand that the primary purpose of this effort was to provide information which could be used to revise and improve the instructional system. Consequently, the evaluation process described here was intended to facilitate instructional decisions, not to yield generalizable data. This evaluation, therefore, necessarily took place within an instructional context, not a research context. Throughout this paper it is helpful to recognize that the data needs of an instructional system--and program managers--are quite different from those of the research study--and researchers. The data need in this instance was reasonably valid, reliable and "inexpensive" information about the twelve instructional system characteristics which have been specified.

Data Needs

Given its task, the sub-group began to consider the various types of data which might be used to determine the extent to which each of the twelve characteristics was present in the program. Each sub-group member, acting independently, took each of the twelve characteristics and generated sets of indicators which one might expect to find if each of those instructional system characteristics were fully in operation. The resulting lists of indicators were discussed; the merit of each indicator was debated. A consensus process was used by the sub-group to generate a single, rather extensive, list of agreed-upon indicators for each characteristic.

The next step was to determine which of those indicators would be used to determine the extent to which the characteristics were present in the program. Two criteria were used in making this decision: (1) which of the indicators appeared to be most important; that is, which seemed to capture the essence of the characteristic; and (2) for which of the indicators was it reasonably feasible to collect data during the spring 1977 semester. Through lengthy discussion, the three members of the sub-group reached consensus and subsequently presented their recommendations to the total task force. The task force considered the recommendations and suggested some modifications; these suggestions were accepted. In this way there was agreement about the indicators to be used.

Data Collection Procedures

The indicators about which data were to be collected were of two types: (1) perception; and (2) documentation. It was determined that perceptual data concerning the twelve characteristics should be collected from students, instructors, supervisors, and supervising teachers associated with the program. Several types of data gathering procedures were considered to be appropriate for this purpose. However, it was decided that these data, given feasibility considerations, would best be collected through the use of questionnaires--a reasonably good means of collecting perceptual data. This decision led to the construction of four questionnaires, one for each of the four groups to be sampled. Additionally, it was determined that data concerning the characteristics might be gained through analyses of various program documents, policy statements, and instructional materials. It was decided that the use of checklists might be an effective means for this purpose, since a checklist can be designed to provide fairly easily obtained, objective information as to whether or not a given characteristic is present.

Questionnaires were mailed to all supervising teachers and student teachers toward the end of the spring semester. At about the same time, questionnaires for the instructors and university supervisors were distributed by instructors in selected classes so that a copy might reach every student enrolled in the PTPP. Questionnaires were returned by way of the same channels.

In order to obtain the program documents which were to be analyzed, the sub-group provided the associate dean for undergraduate studies with a list of each of those documents which was needed. The associate dean, in turn, requested various individuals with PTPP responsibilities to provide the specified documentation. These documents were to be sent to the associate dean, who passed them on to the sub-group for analysis.

Some Findings

Before briefly describing a few of the study's findings, it is perhaps wise to consider several of the factors which may limit them. First, the findings reported here are valid to the extent that the indicators chosen to represent the program characteristics do in fact reflect those characteristics do in fact reflect those characteristics. Second, the accuracy of the findings is influenced by the extent to which the data chosen for analysis represent the selected indicators. Third, the findings and conclusions are limited by the fact that the return rate on the questionnaires and the requested documentation was only moderately high. Fourth, the findings are limited by the reliability and validity of both the questionnaires and the analysis of the documents; no attempt was made to prepare statistical estimates of their validity and reliability. And fifth, the findings are limited to the extent that questionnaire responses and program documentation were accurate and honest.

Space precludes reporting more than just a few of the findings from the questionnaire data and the document analyses. Several of the more interesting items are described here to provide examples of the kinds of data obtained and the manner in which they were treated.

Questionnaire Data

The first twelve items on each of the four questionnaires which were administered to the students, instructors, supervisors, and supervising teachers in this study consisted of statements which presented each of the program's assumptions concerning an effective instructional system. Using a nine-point scale, with one being "strongly disagree" and nine being "strongly agree," respondents were asked to indicate the degree to which they felt that the program had been consistent with each assumption. For example, the first statement on each of the four questionnaires was concerned about the extent to which respondents perceived the program to be competency-based.

The statement to which respondents were to react said: "The program is competency-based; that is, the competencies students are asked to demonstrate are made explicit, the criteria applied in assessing the competencies of students are made explicit, and students are held accountable for meeting those criteria."

Table 1 presents the results obtained from students, instructors, supervisors and supervising teachers for those first twelve items. Mean score data suggest that all four groups agreed that the program had operated in a manner which was consistent with the stated assumptions. Mean scores ranged from a low of 5.13 to a high of 7.85. Student mean scores ranged from 5.99 (personalized) to 7.29 (campus-centered and field-oriented), with the other ten scores falling between 6.10 and 7.04, a rather small range.

Instructor mean scores ranged from 5.13 (personalized) to 7.64 (campus/field); overall, they were a bit lower than student mean scores, with five

Table 1

Summary of Student, Instructor, Supervisor, and Supervising Teacher Questionnaire Data for All Twelve Program Characteristics, Variables 1 through 12. Items 1 through 12

Program Characteristic	Students			Instructors			Supervisors			Supervising Teachers		
	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD
Competency Based	510	6.55	1.99	44	6.55	1.66	23	5.61	1.88	213	6.33	2.31
Campus/Field	503	7.29	1.70	44	7.64	1.66	23	6.65	2.08	197	7.26	1.88
Role-Model Based	493	6.24	1.96	42	5.88	1.99	21	5.38	2.22	197	6.80	1.94
Criterion-Referenced	496	6.49	2.11	43	6.02	2.45	23	6.96	2.16	206	6.77	2.20
Pluralistic	503	7.04	2.10	43	6.05	2.25	23	6.91	2.15	206	7.42	1.92
Humanized	509	6.92	2.04	44	6.41	2.43	23	7.47	1.90	212	7.85	1.72
Personalized	507	5.99	2.37	45	5.13	2.70	23	5.78	2.65	196	6.95	2.21
Modularized	501	6.98	1.95	45	6.71	1.97	22	6.23	2.14	190	6.52	2.14
Multi-Institutional	510	6.98	1.80	45	5.96	2.54	23	6.39	2.54	212	7.73	1.79
Systemized	508	6.33	2.05	45	6.11	2.35	23	6.65	2.04	197	7.17	1.83
Regenerative	491	6.35	2.09	45	5.64	2.59	22	6.77	2.07	191	7.35	1.77
Single System	505	6.10	2.03	45	5.24	2.52	22	5.73	2.12	193	6.90	1.99

scores ranging from 5.13 to 5.96, and another six scores ranging from 6.02 to 6.71. Supervisor mean scores ranged from 5.38 (role-model based) to 7.47 (humanized); four of these ranged from 5.38 to 5.78, while seven ranged from 6.23 to 6.96. Overall, the mean scores obtained from supervising teachers were the highest; they ranged from 6.33 to 7.85.

These results were encouraging. While the data suggest that there was much room for improvement, they also indicated progress. The data from university personnel--instructors and supervisors, those most closely involved in the program's development--indicated that they agreed the program was operationalized in a way which was consistent with the specified assumptions. However, they were the groups which perceived the greatest lack of congruence between program intent and program reality. They seem to have said that "we're getting there but we're not there yet." Student data suggest that students also agreed that the program was reasonably consistent with the assumptions on which it was based. However, the data also suggest areas needing improvement. On the other hand, data from supervising teachers seem to indicate that they felt the program was well on its way to "getting where it wanted to go." One might speculate that their views are colored by their experiences with other teacher education programs, including those in which they were undergraduates.

In summary, then, the data from the first twelve items on each of the four questionnaires, as presented in Table 1, suggest that while progress had been made in operationalizing the instructional system characteristics, there existed much room for improvement.

In an attempt to gain a greater understanding of student perceptions, student questionnaire data were analyzed according to the respondent's certification program (elementary education; secondary education; art education; health and physical education; or music education), and his or her standing (Phase I, II, III, or IV). Two examples of these data are presented in Tables 2 and 3. Table 2 presents the results for Item 1, the item which focused on the students' views regarding the extent to which the program was competency-based. Table 3 presents the results for Item 8, which was concerned with the degree to which the program was modularized.

Table 2

Summary Data from Student Questionnaire Item 1. The program is competency based; that is, the competencies I was asked to demonstrate were made explicit, the criteria applied in assessing my competencies were made explicit, and I was held accountable for meeting those criteria.

Students	N	Mean	SD
Elementary Education Majors			
Phase I	40	6.20	2.10
Phase II	58	6.28	2.25
Phase III	69	6.83	1.73
Phase IV	84	6.51	2.03
Secondary Education Majors			
Phase I	34	6.77	1.99
Phase II	36	6.81	2.12
Phase III	45	6.42	1.73
Art Education Majors			
Phase I	5	7.20	2.49
Phase II	10	7.10	1.52
Phase III	10	6.60	2.27
Health and Physical Education Majors			
Phase I	16	7.31	1.85
Phase II	21	6.76	1.51
Phase III	28	6.21	2.23
Music Education Majors			
Phase I	5	7.00	1.00
Phase II	9	6.67	2.00
Phase III	15	6.73	1.79
Total	510	6.55	1.99

Table 3

Summary Data from Student Questionnaire Item 8. The program is modularized; that is, my instruction was accomplished through the utilization of instructional modules; instructional modules are sets of learning activities which are intended to facilitate my acquisition and demonstration of a particular competency or set of competencies

Students	N	Mean	SD
Elementary Education Majors			
Phase I	40	6.90	1.92
Phase II	59	6.61	2.36
Phase III	68	7.22	1.66
Phase IV	85	7.19	1.41
Secondary Education Majors			
Phase I	34	7.79	1.51
Phase II	35	7.83	1.58
Phase III	45	6.98	2.04
Art Education Majors			
Phase I	4	6.00	1.41
Phase II	8	5.50	3.07
Phase III	9	6.00	1.58
Health and Physical Education Majors			
Phase I	16	7.44	1.37
Phase II	18	6.89	1.97
Phase III	28	6.36	2.30
Music Education Majors			
Phase I	5	7.20	1.30
Phase II	9	6.56	2.07
Phase III	14	5.71	2.64
Total	501	6.98	1.95

The mean scores presented in Table 2 range from 6.20 to 7.31. When subjected to analysis of variance procedures, no statistically significant differences were found between groups. It was concluded that the groups did not differ significantly with regard to their perception of the degree to which the program was competency-based. On the other hand, statistically significant differences were found when the data presented in Table 3 were analyzed using analysis of variance and planned comparison procedures. These mean scores ranged from 5.50 to 7.83. In general, the results indicated the perceptions of art education majors differed from those of secondary education majors. It seemed reasonable to conclude that their experiences differed with regard to the utilization of instructional modules. This suggested that there might be benefit in further examining the program so that these differences might be explained.

The first twelve items of the questionnaire provided data relevant to broad areas. Data from other questionnaire items--and the document analyses--revealed more specific problems. While the first twelve items focused on the total program, the other items tended to focus on specific aspects of the program. For example, on the student questionnaire four groups of items, each consisting of four items for a total of sixteen were concerned with the extent to which four courses or groups of courses were competency-based. An additional four items focused on the degree to which each of those four courses or groups of courses were modularized. The data from these items are presented in Tables 4 and 5.

Data reported in Table 4 indicate that one of the four courses or groups of courses was perceived by students to have been considerably less competency-based than the other three. Mean scores for FED 361: Foundations of Education, ranged from 5.15 to 5.76, while scores for the others (C&I 362: Introduction to the Profession of Teaching; C&I 333: Generic Teaching Competencies; C&I 335: Multicultural Education; and the various methods courses) ranged from 6.70 to 7.20. Likewise, as indicated by the data in Table 5, that same course was viewed as less modularized (mean score of 3.57) than the other three (mean scores of 7.09, 7.17, and 6.22). These findings generated great concern regarding the structure of the foundations of education courses.

Table 4

Summary Data from Student Questionnaire for Variable 1, Items 36 through 51

Item	Course	N	Mean	SD
36	FED 361	391	5.27	2.87
40	FED 361	393	5.15	2.85
44	FED 361	386	5.31	2.85
48	FED 361	385	5.77	2.80
37	C&I 362-333	476	7.20	2.06
41	C&I 362-333	480	6.00	2.08
45	C&I 362-333	480	6.86	2.20
49	C&I 362-333	462	7.08	2.11
38	C&I 335	210	7.04	2.05
42	C&I 335	205	6.71	2.31
46	C&I 335	208	6.75	2.32
50	C&I 335	201	6.90	2.21
39	Methods	342	6.94	2.09
43	Methods	337	6.77	2.09
47	Methods	330	6.84	2.01
51	Methods	323	6.88	2.04

Table 5

Summary Data from Student Questionnaire for Variable 8, Items 114 through 117

Item	Course	N	Mean	SD
114	FED 361	360	3.57	2.92
115	C&I 362-333	454	7.09	2.20
116	C&I 335	194	7.17	2.05
117	Methods	310	6.22	2.57

Document Analyses

In addition to the collection and analysis of perceptual data, the evaluation--as noted earlier--included an analysis of various pertinent program documents. As with the questionnaire data, the intent was to look for evidence that the program characteristics which had been specified as assumptions had been operationalized. The sorts of documents which were requested included:

1. Descriptions of all instructional modules used in PTPP. These were examined for evidence of competency-based instruction, modularized instruction, and criterion-referenced assessment, for example.
2. Descriptions of faculty, program area, and department grading policies and copies of all course grades given to students during the fall, 1976, semester. These were examined for evidence of criterion-referenced assessment procedures and grading, for example.
3. Program policy statements. These contained descriptions of formally adopted goals, assumptions, and required student competencies. These were examined for evidence of a number of the characteristics.

Because of the module descriptions analyses represented the most time consuming of the document analyses aspect of the evaluation and because they yielded some of the most interesting information, the findings obtained from those analyses are highlighted here. Instructors were asked to submit descriptions of the instructional modules used in each course. Each set of materials was examined to determine whether or not the agreed-upon instructional module components were present. That is, each was analyzed to determine whether it contained a rationale, objectives, prerequisites, pre-assessment procedures, learning alternatives, post-assessment procedures, and remediation. Further, a checklist was used to rate each component of each module according to whether it: (1) was present and met the specified criteria; (2) was present but did not meet the specified criteria; and (3) was not present.

Of the thirty-six courses examined, only seventeen were found to utilize instructional modules in the delivery of instruction, while nineteen either submitted no instructional materials or presented materials which were not organized in a modular format. For those seventeen courses which were modularized, a total of 178 module descriptions were examined. It was found that in only six courses were more than two-thirds of the components present and adequate according to the criteria. In descending order, the most often absent components were: remediation, prerequisites, pre-assessment, and post-assessment. Rationales were usually present in the module descriptions but many did not meet the specified criteria. The analysis led to the conclusion that less than half of the PTPP courses were modularized and of those which were, the modules were often deficient in terms of the components they were to contain. Table 6 presents the data for C&I 362: Introduction to the Profession of Teaching, a course for which the findings were rather typical.

Data Utilization

In the beginning of this paper the major purpose of this evaluation effort was described as a process that was to provide information on which to base decisions relevant to program improvement. Thus far the discussion has focused on data collection and analysis. In this section, a brief example of how the data have been used to date is provided.

Findings from the module descriptions analyses were made available to program managers in June, 1977. They used the findings to begin to identify priority areas needing improvement. As a result of their deliberations, a task force was formed in the summer of 1977, and charged with responsibility for revising instructional modules for C&I 362 and 333. The products of that task force, thirty-nine modules, were designed so as to ensure that each contained all of the components that instructional modules were expected to contain, and to ensure that each contained all of the components that instructional modules were expected to contain, and to ensure that all components met the criteria which had been established. An analysis of the thirty-nine instructional module descriptions produced by the task

Table 6

An Analysis of the Instructional Modules: C&I 362. Introduction to the Profession of Teaching

Module	Rationale	Objectives	Prerequisites	Pre-Assessment	Learning Alternatives	Post-Assessment	Remediation
1	2	2	0	0	2	0	0
2	2	2	0	0	2	0	0
3	2	2	0	0	2	0	0
4	2	2	0	0	2	0	0
5	2	2	0	0	2	0	0
6	1	2	2	2	2	2	0
7	1	2	2	2	2	2	2
8	1	2	2	2	2	2	2
9	0	0	0	0	0	0	0

Code: 0 = Component not present.
1 = Component present; did not meet criteria.
2 = Component present; met criteria.

force confirmed that to be the case. In the fall of 1977, those modules were utilized with all C&I 362 and C&I 333 students. During the semester, student and instructor opinions were sampled through the use of feedback forms for each module. Data from students and instructors were very positive. Few weaknesses were identified; recommended modifications were relatively minor. Student questionnaire data from the fall semester indicated some improvement with regard to student perceptions of the extent to which the program was modularized. For example, on the questionnaire item dealing with modularization, a mean score of 6.90 had been obtained for the forty elementary education students who took C&I 362 during the spring, while a mean score of 7.17 was obtained for the eighty-two who took it in the fall. As additional data become available to program managers and the faculty, continued efforts will be made to improve the program. Such efforts are based on the instructional system assumption concerned with program regeneration. It has been assumed that the program will be more effective if it operates as an open system, capable of continuous revision on the basis of constructive data supplied by sound formative and summative evaluation procedures.

Future Steps

The task force plans to continue the PTPP evaluation efforts that were initiated in 1976. The task force will continue to collect data relevant to instructional system characteristics, student characteristics, and student satisfaction and success. The evaluation plan will be expanded to examine the operations of the learning resource center and to provide more specific information on particular components of the program. At this point in time, serious consideration is being given to collecting data every other year, rather than yearly, because of cost factors and because this would allow program components time to make the modifications called for by findings from the evaluation information. Moving to an every-other-year approach might also avoid the potential problem of alienating students and faculty who might resent spending the time which data collection and program modification requires.

How do we evaluate our evaluation efforts? Generally, we are pleased with the kinds of data we collected on this first attempt. We are displeased with the tremendous time lag between when some of the data were collected--spring, 1977--and when they were analyzed--winter, 1977. This time lag caused us to fail to meet one of our goals of the evaluation, that is, to present data that are timely for decision-making purposes. We encountered a plethora of problems related to the coding of data, key punching, computer programming, and computer operations.

Some of these problems were of our own making and were largely due to our inexperience, while others resulted from circumstances beyond our control. Regardless of the reason, we do not feel that we can award ourselves a passing grade on the criterion of timeliness. The experience we have gained, however, should result in tremendously improved efficiency in the data collection and analysis processes. In summary, we learned a lot, we liked what we did, we found it worthwhile, and we are going to do it again.

PROFESSIONAL TEACHER PREPARATION PROGRAM EFFECTIVENESS STUDIES: 1976-1977
AFFECTIVE TESTING

Howard Jones, Robert Randall
University of Houston

PROFESSIONAL TEACHER PREPARATION PROGRAM EFFECTIVENESS STUDIES:
1976-1977 AFFECTIVE TESTING

Howard Jones and Robert Randall
University of Houston

During spring, 1977, students in all phases of the undergraduate Professional Teacher Preparation Program at the University of Houston were administered four instruments in order to assess self-image, attitudes toward teaching, motivations and cognitive styles. This report documents some of the results of this assessment effort.

Teacher education research literature is replete with a number of so-called presage-process or presage-product studies. The goal of these studies is to determine if there are any personality, attitudinal, or valuing characteristics associated with more effective teachers. Most studies have not been useful to program builders. While Ryans (1960) and Sandefur (1976) have identified some personal characteristics (such as warmth), most studies have noted that effective teachers come in all styles, demonstrating a variety of affective dimensions.

So why administer affective instruments to undergraduates? The answer to this question focuses on program assessment rather than student assessment and has three parts. First, if the PTPP program is useful, then students' personality characteristics should not be negatively effected as a result of going through the program. For example, a student's self-image should be no lower at the end of the program than it was when he or she began. Second, there is the issue of student satisfaction with the program. Research on cognitive styles of learners has indicated that a good predictor of student satisfaction is the overlap of student cognitive style with that of faculty members. There are some characteristics of the Houston program (self pacing, modularization, etc.) that may not be as agreeable to some students as it is to others. If it were possible to identify whether there are students with some personality/attitidinal/cognitive style characteristics who are not satisfied with the Houston program, these data could be used to plan alternative efforts with students. Third, which is probably of interest to researchers, there is the issue of affective growth of students in teacher education programs. Fuller's (1969) research, as well as that reported by Coates and Thoreson (1976), notes what all teachers know--that teaching is an anxiety producing effort. A study of possible effective changes in students as they approach student teaching could add much to the literature.

The Plan

With these issues in mind, the PTPP Research Committee identified four instruments to be administered to students during the spring of 1977. In selecting the instruments the following four-step rationale was used:

1. The selected instruments would have a reasonably high face validity for faculty and students.

Affective instruments have been administered to students for several years in the PTPP program. Several instruments have been useful in aiding students; but, of these, one or two have not had high face validity. In short, they were excellent psychological tests, but many undergraduates did not consider them beneficial.

2. The selected instruments would have local norms.

Coupled with the lack of face validity, some instruments are normed on other populations rather than teachers, or they are normed on teachers who have different characteristics from students at the University of Houston. For this reason, it was decided that instruments found useful in the past at Houston and having UH norms, or normed on a similar population, would be used.

3. The selected instruments would have outcomes having some degree of overlap with existing program goals.

Since a number of objectives in the PTPP program relate to self-awareness, it appeared that affective instruments which assess objectives would be useful in programmatic assessment.

4. The selected instruments could be scored and the results interpreted easily by both students and faculty.

In order to side-step another difficulty with past affective assessments, it was decided that faculty members in the program would provide feedback to students. For this reason, the instruments selected were relatively easy to interpret. Instruments with computer scoring capability were also selected.

The Selected Instruments

Minnesota Teacher Attitude Inventory (MTAI)

This instrument has been used with more success than any other single identifier of presage characteristics. Initially, the instrument had a single score permitting a global "g" factor identification of teacher attitudes. However, work by Fruchter and Yee (1971), as well as efforts by Shores (1977), permits a breaking of the one score into five separate factors:

a. Children's Tendencies--does the teacher consider students as needing control or as being self-directed?

b. Pupil Cooperation--is the pupil considered as cooperative or as needing coercion to learn?

c. Handling Pupil Behavior--must the teacher control all pupil behavior? Should the teacher be strict or sympathetic and understanding?

d. Pupil's Independence in Learning--is there a high or low need for a teacher to keep order in a classroom?

e. Pupil's Acquiescence--do children relate to teachers in an antagonistic or acquiescent fashion?

In addition, based on the work of John Bell (1977), items representing two other scales, Traditionalism and Progressivism, were included in the final instrument. These scales permit a study of the way prospective teachers view the roles they play as teachers and how this view relates to subject content, (traditionalism) or on learner needs (progressivism).

Adjective Self Description (ASD)

This assessment permitted a study of students on the following scales:

a. Social Attitude (cold-warm).

b. Social Behavior (courteous-annoying).
c. Performance Habits (careless-efficient).
d. Social Orientation (outgoing-reserved).
e. Emotional Stability (serene-anxious).
f. Ideological Orientation (practical-idealistic)
g. Appearance and Charm (plain-attractive).

Obviously, the focus of this assessment is on self-perception. The assessment itself consists of 56 objectives to which the student responds on a 1-5 Likert scale. The items are actual extensions of the seven factors above. Students in the past have found the ASD to have high face validity and to be easily interpreted.

Work Motivation Inventory (WMI)

As an attempt to assess students' motivations according to an agreed upon paradigm, the Work Motivation Inventory (WMI) was selected for use. The WMI was developed by Hall and Williams (1973) to tap motivations of businessmen in the light of Maslow's needs hierarchy--a hierarchy cognitively explored in-depth by students in all PTPP programs (Maslow, 1954). Thus, the use of an instrument with a recognized theoretical framework was seen as an advantage. The disadvantages of the instrument were two-fold: (a) difficulty of administration (the items are unusual, and care must be made in giving directions to students); and (b) norms (the norms reported by Hall and Williams concern businessmen, who probably have different needs than teacher education students). However, a study of past administration efforts permitted a renorming using UH students, and the instrument was selected for use.

The WMI provides data for students on the following scales:

a. Basic Needs
b. Safety Needs
c. Belonging Needs
d. Ego-Status Needs
e. Actualization Needs.

Cognitive Style Mapping (CSM)

The three instruments identified above focus on attitudes, self-perceptions and motivations. The last instrument, Cognitive Style Mapping, was selected to permit a focus on student identification of patterns around them. The concept "cognitive style" has been explored by Witkin (1977), and Kagan (1963), among others. The term itself is open to a number of interpretations. Hill (1970), at Oakland Community College, has developed an instrument with 28 scales, permitting the student to identify his preferences for taking in information. The Cognitive Style map developed by Hill is based on a number of theoretical constructs.

In initial exploration efforts at UH, five of the scales (the sensory scales) were not found to be very useful. Most UH students look remarkably alike on these scales. In 1975, with permission from Hill, the original instrument was modified, reducing the level of readability and focusing on the remaining three scales.

While the instrument may appear to be more research-oriented than feedback-oriented, the latter is the rationale Hill used for its development. At Oakland (as well as other colleges), Hill et al. prescribed learning activities for students based on the results of the CSM data. For example, students scoring high on some scales have been found to prefer to work alone; other students prefer to work in groups. The key word in this last sentence is prefer. The selection of the CSM instrument for use in this research effort was in order to identify whether there is a set of CSM preferences that showed up in students who are more (or less) satisfied with the PTPP program.

Administration of Instruments

During January and February of 1977, Jones and Randall met with faculty members in all phases of all programs, explaining the instrumentation and the detailed testing requirements. Faculty members were to administer the instruments on a voluntary basis. Students not wishing to be involved were not to be forced to take the instruments. It was anticipated that this would lower the generalizability of the test results somewhat, but the acceptance of the assessments was considered important in obtaining valid responses in this pilot effort.

Instruments were administered by individual faculty members during

February and March, 1977. Tests were scored and returned to faculty as quickly as possible for interpretation sessions.

A training program for faculty test interpretation was developed and implemented by faculty at the University of Houston. Faculty were provided detailed notes and overhead projections describing each instrument. For interpretation of one instrument, Cognitive Style Mapping, a self-analysis packet was developed and distributed to students along with their cognitive style scores. For copies of these instruments write the authors at the address noted in the "List of Contributing Authors."

Data Analysis

The evaluation plan did not call for a comparison of students within phases nor for a comparison of faculty members within or across phases. For this reason, after individual students had received their feedback, all data were combined into the following groups:

Elementary	Phase 1	
Elementary	Phase 2	
Elementary	Phase 3	
Elementary	Phase 4	
Music Ed.	Phase 1	
Music Ed.	Phase 2*	
HPE	Phase 1	(Health & Physical Ed. majors)
HPE	Phase 2	
HPE	Phase 3	
SED	Phase 1	(Secondary Ed. majors)
SED	Phase 2	
SED	Phase 3	
Bil Ed. Elementary	Phase 1	(Elementary Ed. majors also seeking bilingual endorsement)
Bil Ed. Elementary	Phase 2	
Transfer Students Elementary	Phase 2	(All Elementary Ed. majors transferring from other institutions are required to take this course)

*No data were returned from Music Education Phase 3 students.

Art Ed.	Phase 1
Art Ed.	Phase 2
Art Ed.	Phase 3

Summary of Minnesota Teacher Attitude Inventory Results

There appears to be more similiarities between programs and students than there are differences in this instrument. Results indicate that students in the program think that children are self-directed and cooperative. The results also reflect attitudes of sympathy toward handling pupils and helping students with their own interests. Students also reflect the attitude that pupils do work to please the teacher. While students are somewhat traditional in their view of the roles of schools, they reveal a progressive attitude toward learner-centered interests in teaching.

Summary of Adjective Self-Description

Based on the results of this analysis, it appears that most teacher education students see themselves as warm, courteous, efficient, somewhat outgoing, somewhat anxious, idealistic and attractive. While there are some significant differences among the population groups, there appears to be no one scale that would separate groups from each other.

Summary of Work Motivation Inventory

There are few differences noted among students in programs and phases on the Work Motivation Inventory. It has been pointed out that UH teacher education students are motivated by belonging needs as well as ego-status and actualization needs. The motivations of UH students are clearly different from the motivations of businessmen. (For additional information on the differences between students and businessmen write the authors at the address given in the "List of Contributing Authors.") There is, however, a remarkable consistency across programs in terms of motivations of students in the PTPP program.

Summary of Cognitive Style Mapping

Figure 1 summarizes the findings of this part of the study. To the right of each identified scale is a rectangle representing the interquartile range for all students in the program. The vertical line in each box represents the mean. The heavy vertical line through the entire graph would indicate the point above which students would be identified as a "major" by Hill (1970).

With the exception of T(AQ), Q(CH), A and F, our students appear to be "majors" in the scales. This might be a function of the fact that teacher education students at UH are at least juniors. Those students who have succeeded in at least two prerequisite years in college were able to cope with a variety of teaching-learning environments.

As noted earlier, the combination of student scores on scales prohibits the use of the Cognitive Style Mapping instrument as a diagnosis-prescriptive tool -- for which it was originally designed. However, in looking at the combined scores for the 559 students who took the instrument, it would appear that the following predictions might have some validity:

1. If students are to acquire numerical data, it best be done through written rather than oral (taped, lectured) presentations.

2. Our students appear to prefer written verbal material to lecture situations.

3. Since a significant portion of students are minors, and not majors, on the histronic scales, role-playing and video-tape sessions should be less than comforting for students.

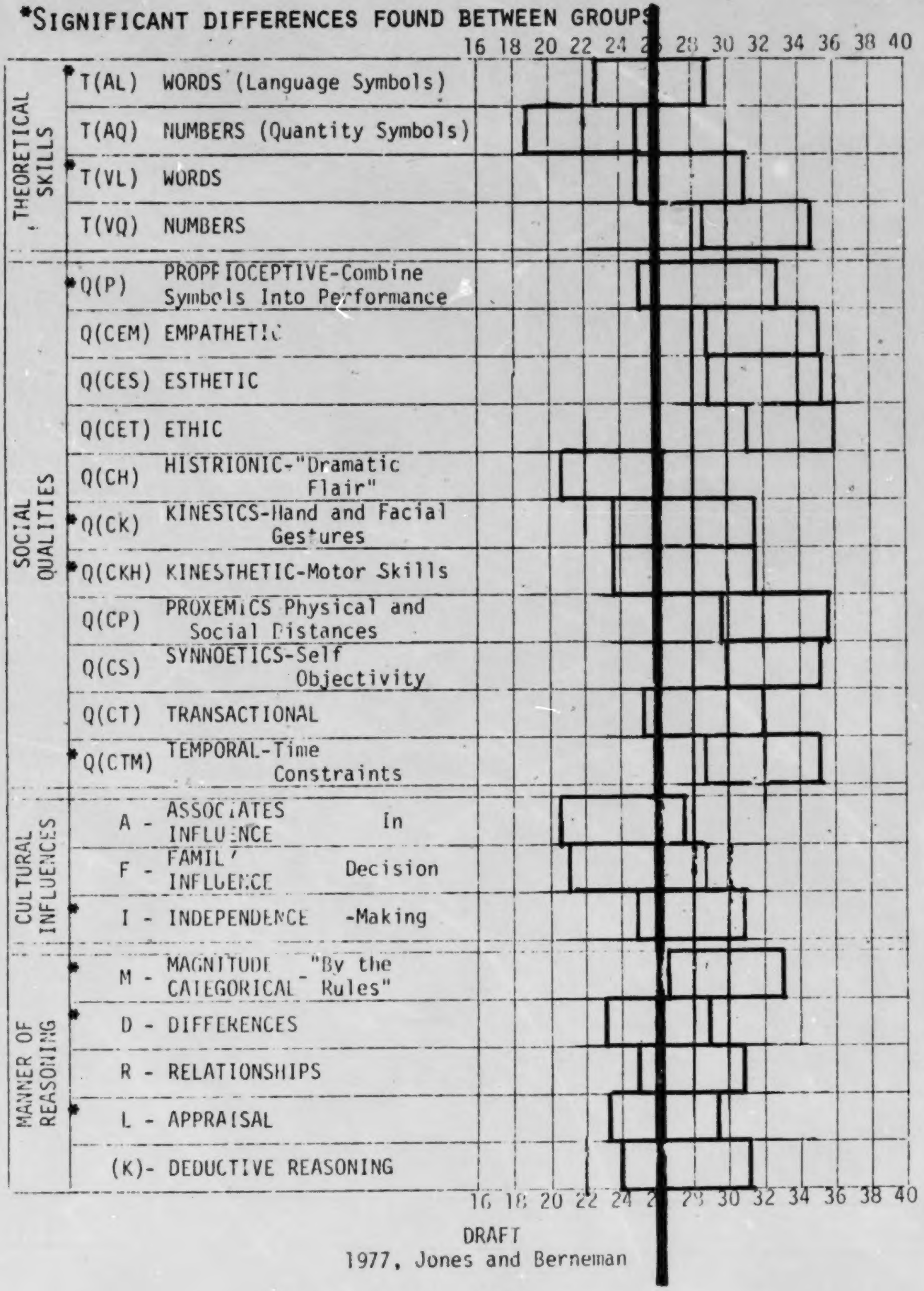
4. Our students are quite time conscious. This might reflect the fact that many have jobs and families in addition to their responsibilities as students. The identification of realistic time estimates from faculty members would appear to be a critical effort.

5. With respect to Cultural Influences it appears that our students are somewhat more independent than influenced by associates (A) or family (F). Based on this it might be recommended that our current practice of providing learning options for students (in which they place their own structure)

FIGURE 1
COGNITIVE STYLE MAP

Name N = 559 SPRING, 1977 Date

Student No. Classification FR SO JR SR GR


*SIGNIFICANT DIFFERENCES FOUND BETWEEN GROUPS

DRAFT
1977, Jones and Berneman

be continued.

6. The Manner of Reasoning of Students appears to indicate that most students are capable of all types of logic.

References

Bell, J. Stability of the factor structure of an attitude inventory. Short form of the Minnesota Teacher Psychology in the Schools, 1977, Vol. 14, 169-171.

Coates, T., & Thoreson, C. Teacher anxieties: A review with recommendations. *Review of Educational Research*, Spring, 1976, *46*(2), 159-184.

Fruchter, B., & Yee, A. Factor content of the Minnesota teacher attitude inventory. *American Educational Research Journal*, 1971, *8*, 119-133.

Fuller, F. F. Concerns of teachers: A developmental conceptualization. *American Educational Research Journal*, 1969, *6*(2), 207-226.

Hall, J., & Williams, M. *Work motivation inventory*. Conroe, Texas: Telemetrics, 1973.

Hill, J. E. *Cognitive style as an educational science*. Bloomfield Hills, Michigan: Oakland Community College Press, 1970.

Kagan, J., Moss, H. A., & Siegel, I. E. Psychological significance of styles of conceptualization. In J. C. Wright and J. Kagan (Eds.), *Basic cognitive processes in children*. *Monographs of the Society for Research in Child Development*, 1963, *28*(2, Serial No. 86), 73-112.

Maslow, A. *Personality and motivation*. New York: Harper & Row, 1954.

Sandefur, J. T. Research supported generalizations on teaching and teachers. Unpublished manuscript, Western Kentucky University, 1976.

Shores, J. Houston, Personal communication, 1977.

Ryans, D. G. Characteristics of teachers. Washington: American Council on Education, 1960.

Witkin, H. A., Moore, C. A., Goodenough, D. R., & Cox, P. W. Field-dependent and field-independent cognitive styles and their educational implications. *Review of Educational Research*, 1977, 1-64.

PTPP SUCCESS STUDY OF STUDENT TEACHERS

Howard Jones, Robert Randall
University of Houston

PTPP SUCCESS STUDY OF STUDENT TEACHERS

Howard Jones and Robert Randall
University of Houston

PTPP program developers at the University of Houston face a dilemma when determining the success of their graduates. The graduates, both those completing student teaching and inservice teachers, should have a positive impact on their pupils. By assessing this impact using standardized or other tests, an indicator of teacher success should be evident. This "accountability" model is considerably easier to discuss than it is to implement, however. The variance of school settings, class sizes, available teaching aids, pupil aptitudes and motivations all work to deter valid accountability evaluations.

A second possibility of evaluating PTPP graduates--student teachers and inservice teachers--is to collect data on the teachers as they teach. This step calls for classroom observations and the use of high or low inference instruments to identify the teachers' processes with pupils. Difficulties with the implementation of this effort are also evident. The time needed for actual classroom data collection is only one problem. Furthermore, the PTPP program at the University of Houston has placed a high positive value on the use of clinical supervisory modes. In such modes, supervisors work with student teachers and School-Based Teacher Educators in assessing, in a formative mode, the processes of the student teachers. By implementing what could be perceived as a "big brother" approach to data collection, it was felt that much of the positive affect associated with the clinical supervision mode would be damaged.

To collect data on the success of students in the PTPP program, it was decided that a high-inference rating scale would be used. The scale would be administered to student teachers and their supervisors. In addition, the School-Based Teacher Educator (SBTE) in whose classroom the student teacher worked would be asked to rate the student teacher on the scale items.

In April, 1977, near the completion of the student teaching experience, all student teachers, their supervisors and School-Based Teacher Educators were administered a rating scale to assess the success of the student teacher.

The instrument consisted of a series of 1-5 Likert scales on which the rater identified his or her ratings. The items reflected 15 of the 16 generic teaching competencies of the undergraduate program--competencies familiar to _all_ students in _all_ programs.*

Because a few competencies are quite broad, these competencies were assessed in two or more stems. For example, Competency Eight (Promotes Effective Patterns of Communication), was subdivided into: Communicates with Students; Communicates with other Teachers; Communicates with Supervisors; Communicates with Parents. Furthermore, to identify some level of perceived importance of each of the competencies, raters were asked to identify (again on a 1-5 scale) the importance that they associated with the statement. In cases where raters felt they could not respond to an item, they were instructed to indicate "Can't Rate."

In essence, the raters (students, supervisors, School-Based Teacher Educators) both indicated their perception of student teachers' demonstration of program competencies and identified their personal feelings about the value of each competency statement.

Anonymous responses were received from 191 student teachers, 22 University supervisors of the student teachers, and 223 School-Based Teacher Educators. Data analysis of these responses are provided in the next section.

Data Analysis - Student Teacher Ratings

COMPETENCY ONE: IDENTIFIES LEARNER'S EMOTIONAL, SOCIAL, PHYSICAL AND INTELLECTUAL NEEDS.

Two stems were developed for assessing teacher/supervisor ratings of this competency:

---

*The sixteenth competency reflecting students' incorporating career education concepts into the classroom was not assessed, since it recently added to the program. Student teachers would have had no planned experiences in this competency.

Stem 1: Apply knowledge of human growth and development in planning for instruction; and

Stem 2: In planning for instruction, apply knowledge about how students learn.

Perceived Importance

As is the case on all the rated items, rater scores can range from 1 (low importance) to 5 (high importance). Tables 1 and 2 note group statistics for all subjects. In addition, they provide a separation of respondents by level (Elementary Education, Secondary Education, etc.). Unfortunately, in returning the completed questionnaires, a significant number of student teachers and SBTE's did not provide data about their levels. The data are categorized in Tables 1, 2, and subsequent tables as "Unknown."

Tables 1 and 2 also reflect the results of one-way analyses of variance designed to test for differences within the groups. For both the student teacher and SBTE's, analyses of variance project possible differences of rating according to level or major.

As can be noted on both Tables 1 and 2, all raters indicated a high value for the competency statements. Only the SBTE average rating of 4.488 on Stem 1 (Table 1) is less than 4.5.

Analysis of variance indicated that there were perceived differences in the value of Stem 1 (Knowledge of human growth and development) by the student teachers. However, in post-hoc analysis, it was determined that this difference was an anomaly caused by the inclusion of "unknown" students in the analysis. Therefore, it was concluded that all student teachers valued the stem highly and there were no significant differences between the ratings of groups of student teachers.

However, there was a significant difference found in SBTE ratings of the importance of Stem 2 (In planning for instruction apply knowledge about how students learn). In post-hoc analysis, it was found that SBTE's in elementary levels rated this stem higher than Music Education SBTE's. However, the reader should note the very high standard deviation found in the analysis of Music Education SBTE's (1.344). This indicates a wide range of opinion with respect to the importance of this stem.

Table 1

Ratings of Perceived Importance of the Stem:

Apply knowledge of human growth and development in planning for instruction

SCHOOL-BASED TEACHER EDUCATORS	217	4.488	0.840	
Elementary	44	4.705	0.553	
Music Educ.	13	4.077	1.605	Obtained F-ratio = 2.24
HPE	19	4.211	1.032	(5,211)
Sec. Educ.	41	4.317	0.934	p = .0520
Art Educ.	12	4.500	1.168	
Unknown	88	4.580	0.601	
UNIVERSITY SUPERVISORS	22	4.727	0.550	
STUDENT TEACHERS	185	4.611	0.692	
Elementary	67	4.716	0.486	
Music Educ.	16	4.437	0.892	Obtained F-ratio = 2.33
HPE	25	4.720	0.458	(5,179)
Sec. Educ.	59	4.407	0.912	p = .044
Art Educ.	8	4.875	0.354	
Unknown	10	4.900	0.316	
	N	Mean	Standard Deviation	Results of One-Way ANOVA within Groups

Table 2

Ratings of Perceived Importance of the Stem:

In planning for instruction apply knowledge about how students learn

	N	Mean	Standard Deviation	Results of One-Way ANOVA within Groups
SCHOOL-BASED TEACHER EDUCATORS	219	4.699	0.7107	
Elementary	44	4.841	0.370	
Music Educ.	13	4.154	1.344	Obtained F-ratio = 3.26
HPE	20	4.450	0.999	(5,213)
Sec. Educ.	41	4.707	0.750	p = 0.0075
Art Educ.	12	4.417	1.164	
Unknown	89	4.798	0.457	
UNIVERSITY SUPERVISORS	22	4.909	0.294	
STUDENT TEACHERS	188	4.617	0.754	
Elementary	68	4.662	0.614	
Music Educ.	16	4.500	1.095	Obtained F-ratio = 0.38
HPE	25	4.520	0.653	(5,182)
Sec. Educ.	61	4.590	0.901	p = 0.8594
Art Educ.	8	4.750	0.463	
Unknown	10	4.800	0.422	

Student Teacher Performance

Next, student teachers, SBTE's and university supervisors were asked to rate the student teachers' skills in implementing the competency with pupils. University supervisors, many of whom had 12-14 students, provided a mean rating for all of their students.

As can be noted on Tables 3 and 4, the ratings of student teacher performance are generally lower than the perceived importance ratings. In fact, the mean university supervisor ratings of the student teachers on both stems was below 4.0. In each case, the student teachers rated themselves higher than the SBTE's, who rated the student teachers higher than the university supervisors. As might be expected, reasonably high standard deviations were found in all cases.

Analysis of variance results indicated only one significant result: student teachers' self ratings on Stem 1 (Apply knowledge of human growth and development...). Post-hoc analysis again indicated that this finding of significance was an anomaly caused by the inclusion of the "Unknown" group in the analysis.

Summary

In general, the ratings from all groups on student teacher performance of Competency One indicate that there is moderate to high demonstration of this competence by student teachers in school settings.

Summary

Tables 1 and 2 report statistical data for each of the Perceived Importance and Student Teacher Success parts of the PTPP Questionnaire. In viewing the entire set of data several elements are evident:

1. There appears to be general agreement with the importance of program competencies. Not one competency stem was rated below 4.124 by any of the three rating groups. There appears to be a general consensus of programmatic thrusts. This is not surprising, considering the effort spent in acquiring SBTE input into the program. However, it is rewarding to see the agreement.

Table 3

Ratings of Student Teacher Performance of the Stem:

Apply knowledge of human growth and development in planning for instruction

	N	Mean	Standard Deviation	Results of One-Way ANOVA within Groups
SCHOOL-BASED TEACHER EDUCATORS	214	4.033	0.895	
Elementary	44	4.250	0.615	
Music Educ.	13	3.923	0.862	Obtained F-ratio = 0.86
HPE	20	4.050	0.944	(5,208)
Sec. Educ.	39	3.897	0.968	p = 0.5074
Art Educ.	11	3.818	1.250	
Unknown	87	4.023	0.927	
UNIVERSITY SUPERVISORS	22	3.727	0.882	
STUDENT TEACHERS	180	4.167	0.729	
Elementary	65	4.338	0.644	
Music Educ.	16	4.125	0.719	Obtained F-ratio = 2.74
HPE	25	4.040	0.611	(5,174)
Sec. Educ.	57	4.018	0.834	p = 0.0209
Art Educ.	7	3.714	0.756	
Unknown	10	4.600	0.516	

Table 4

Ratings of Student Teacher Performance of the Stem:

In planning for instruction apply knowledge about how students learn

	N	Mean	Standard Deviation	Results of One-Way ANOVA within Groups
SCHOOL-BASED TEACHER EDUCATORS	216	4.037	0.969	
Elementary	44	4.182	0.815	
Music Educ.	13	4.000	0.816	Obtained F-ratio = 0.40
HPE	20	3.850	1.089	(5,210)
Sec. Educ.	40	4.000	1.132	p = 0.8507
Art Educ.	10	3.900	1.287	
Unknown	89	4.045	0.928	
UNIVERSITY SUPERVISORS	20	3.682	0.839	
STUDENT TEACHERS	182	4.154	0.757	
Elementary	67	4.269	0.618	
Music Educ.	16	4.312	0.793	Obtained F-ratio = 1.71
HPE	25	4.000	0.816	(5,176)
Sec. Educ.	59	4.051	0.839	p = 0.1337
Art Educ.	7	3.714	0.756	
Unknown	8	4.500	0.756	

2. In all cases of the rating of student performance, student teachers generally rated themselves higher than their corresponding SBTE's and university supervisors. It is difficult to account for this since students were assured to anonymity and were told that their reports would not count on grading, etc.

3. In viewing each of the tables, it appears that Elementary Education SBTE's and student teachers were consistently rating the competencies and student performance higher than their other counterparts. Again, it is difficult to account for this.

Because of the the difficulties in obtaining truly reliable ratings across groups in student performance and perceived importance of the competency statement, an additional step was taken. The mean ratings for each group on each of the scales were ranked within group. For example, SBTE mean ratings of all student teacher performance stems were ranked, giving a ranked order statement of student teacher performance. The highest rated SBTE rating of student teachers was on Stem 21 (Willingness to continually improve as a teacher). The lowest rating was on Stem 2 (Apply knowledge of learning in planning). Thus for the SBTE ratings, Stem 21 would be ranked 1st; Stem 2 would be ranked 22nd.

Table P provides group mean ratings and rankings within groups for each of the stems. In addition, an average ranking, obtained by calculating the mean of ranks for each of the three rating groups, is provided.

Perceived Importance

While all stems were rated quite highly by all groups, the highest three ranked stems are:

No. 21. Demonstrate a willingness to continually improve as a teacher.

No. 11. Communicate with students.

No. 9. Use a variety of teaching skills.

The three lowest ranked stems are:

No. 3. Writes clearly, stating lesson and unit objectives.

No. 8. Design/incorporate materials, etc., which blend with

TABLE P

Perceived Importance					Rating of Student Teacher Performance			
Avg. Rating	SBTEs	University Supervision	Student Teacher	Stem	Student Teacher	University Supervision	SBTEs	Avg. Rating
15.0	4.488 18	4.727 13(T)*	4.611 14(T)*	1. Apply knowledge of human growth and development in planning for instruction.	4.167 21	3.727 20	4.033 18	19.7
9.7	4.699 10	4.909 6(T)*	4.617 13	2. In planning for instruction apply knowledge about how students learn.	4.154 22	3.682 21	4.037 17	20
20.3	4.326 20	4.409 19	4.140 22	3. Write clearly stated lesson and unit objectives.	4.410 12	4.045 11	4.192 11	11.3
16.7	4.571 13	4.454 18	4.426 19	4. Construct understandable and useful lesson and unit plans.	4.484 8	4.182 9	4.252 7	8
19	4.535 16	4.364 20	4.351 21	5. Implement instruction that is based on designed unit and lesson plans.	4.436 11	4.227 7	4.242 8	8.7

*(T) = tie

TABLE P

Perceived Importance					Rating of Student Teacher Performance			
Avg. Rating	SBTEs	University Supervision	Student Teacher	Stem	Student Teacher	University Supervision	SBTEs	Avg. Rating
17.7	4.479 19	4.476 17	4.468 17	6. Construct and use tests and other instruments to monitor student progress.	4.403 13(T)*	3.882 15	3.984 20	19.3
14.3	4.532 17	4.714 15(T)*	4.725 11	7. Report student progress to students and parents.	4.324 17	3.600 22	4.172 15	18
20.3	4.124 22	4.286 21	4.429 19	8. Design and/or incorporate materials, illustrations and examples into lessons that blend with the cultural environments of students.	4.209 20	3.833 18	40.20 19	19
3.7	4.832 4	4.954 3(T)*	4.827 4	9. Use a variety of teaching skills (explains, questions, reinforces, etc.)	4.565 5	4.143 10	4.239 9	8

*(T) = tie

TABLE P

Perceived Importance					Rating of Student Teacher Performance			
Avg. Rating	SBTEs	University Supervision	Student Teacher	Stem	Student Teacher	University Supervision	SBTEs	Avg. Rating
9.7	4.701 9	4.773 10(T)*	4.733 10	10. Use a variety of instructional strategies (gaming, simulations, lecture, small group, individualized instruction, e.g.)	4.393 15	3.950 14	4.185 12	13.7
3	4.878 2	4.909 6(T)*	4.890 1	11. Communicate with students.	4.684 2	4.454 4	4.330 3	3
15.3	4.536 15	4.714 15(T)*	4.607 16	12. Communicate with other teachers.	4.514 7	4.474 3	4.183 13	7.7
12	4.612 12	4.773 10(T)*	4.611 14(T)*	13. Communicate with supervisors.	4.438 10	4.546 2	4.317 4	5.3
11.7	4.543 14	4.762 12	4.738 9	14. Communicate with parents.	4.308 18	4.000 12(T)*	3.819 21	17
21	4.310 21	4.273 22	4.407 20	15. Operate and use when appropriate audio-visual aids (e.g., projectors, recorders).	4.369 16	3.857 17	4.253 6	13
4.7	4.801 5	5.00 1(T)*	4.751 8	16. When necessary, changing plans based on the feedback received from students.	4.478 9	3.864 16	4.182 14	13

*(T) = tie

TABLE P

Perceived Importance					Rating of Student Teacher Performance			
Avg. Rating	SBTEs	University Supervision	Student Teacher	Stem	Student Teacher	University Supervision	SBTEs	Avg. Rating
11	4.757 8	4.727 13(T)*	4.684 12	17. Use a variety of classroom management (discipline) strategies.	4.246 19	3.737 10	3.785 22	20
9	4.695 11	4.857 9	4.762 7	18. React to the needs of students parents, supervisors, other teachers and myself.	4.403 13(T)*	4.190 8	4.155 16	12.3
4	4.842 3	4.909 6(T)*	4.874 3	19. Demonstrate openness and flexibility in teaching.	4.640 4	4.273 6	4.265 5	5
4.7	4.795 6	4.954 3(T)*	4.825 5	20. Exhibit a willingness to change when presented with valid data.	4.652 3	4.409 5	4.344 2	3.3
1.3	4.885 1	5.000 1(T)*	4.881 2	21. Demonstrate a willingness to continually improve as a teacher.	4.787 1	4.591 1	4.547 1	1
6	4.782 7	4.952 5	4.801 6	22. Demonstrate adequate content knowledge.	4.527 6	4.000 12(T)*	4.207 10	9.3

*(T) = tie

multicultural environment of students.

No. 15. Operate and use A-V aids.

Student Teacher Performance

There is a much higher variance of raters' opinions of student teacher success on each of the stems. However, the three items on which student teachers received highest ratings are:

No. 21. Demonstrate a willingness to continually improve as a teacher.

No. 11. Communicate with students.

No. 20. Exhibit a willingness to change when presented with valid data.

Four stems appear to be about equal with respect to the lowest student ratings. These are:

Stem 8. Design/incorporate materials, etc., that blend with multicultural environment of students.

Stem 6. Construct and use tests and other instruments to monitor student progress.

Stem 1. Apply knowledge of human growth and development in planning for instruction.

Stem 17. Use a variety of classroom management (discipline) strategies.

In looking at the ranked elements there are several interesting points:

1. In several instances, PTPP students excell at those competencies considered most important by the raters (Communication with students, willingness to change).
2. In one case, students appear to perform poorly in the design and incorporation of multicultural elements. This was also a low ranking competency statement.

SUMMARIZING THE STUDIES

Shirley M. Hord
Research and Development Center
for Teacher Education, The University of Texas

SUMMARIZING THE STUDIES

Shirley M. Hord
Research and Development Center for Teacher Education
The University of Texas at Austin

There are widespread indications that the training and preparation of teachers is in urgent need of thorough examination and immediate corrective action. In response to this situation, many concerned individuals at a number of institutions have been and are continuing to explore the effects which current methods and practices have on the graduates of their teacher education programs. The accelerated interest in such program evaluation and follow-up studies may be due also to the mandates for meeting accreditation standards or to internal institutional pressures to show evidence of program effectiveness.

The ten papers in this volume report on studies conducted on teacher education programs and graduates from seven institutions. These institutions represent not only large, state-supported universities, but also smaller regional colleges. They are: Oregon College of Education, The Ohio State University, Tennessee Technological University, University of Houston, University of Oregon, Weber State College, and Western Kentucky University. The practical lessons learned from these investigations should prove instructive for other institutions who are planning studies or who are in the process of re-examination of present efforts.

The papers in this collection are organized around certain common elements: the purpose of the study; descriptions of the methodology, including information such as data points, selecting the sample, approaches to design, instrumentation, data collection; data management, analysis and storage are addressed by several authors. Findings are reported by nearly all papers, although the utilization of these findings is not reported by all. Finally, problems encountered and "lessons learned" are shared by all the papers, and costs are cited in several.

Purpose

The studies have a common purpose: to assess outcomes and to use the results to provide for program maintenance and to make program revisions and adaptations. A key assumption in the papers is that program revisions should be based on objective and quantifiable information from a variety of sources, as well as subjective information. It also has been assumed that study data would be used for decision-making by faculty and administrators alike. Several institutions (Oregon College of Education, University of Houston, and Weber) are especially attuned to acquiring data to make judgments about recently revised or implemented competency-based teacher education programs, as opposed to programs with a more traditional approach.

More explicitly, goals of the studies focus on (1) graduates' satisfaction with various elements of the training programs and their views on the utility and sufficiency of program components; (2) judgments of graduates' acquisition of program competencies made by classroom cooperating teachers, college supervisors, building principals, and graduates' peers and students; and (3) the attitudes of users (public school educators) toward the program and the graduates' competencies. Additional objectives include the assessment of the teaching styles used by graduates and the possible effects the school setting (for teaching) might have on each teacher's success.

As a result of their studies, some institutions anticipate that comparative studies among various teacher preparation programs might be undertaken. Another suggestion for utilizing the information provided by the studies is for predicting the effectiveness of teachers. In addition, it is expected that the information gained can enable the pros and cons of various methodologies to be articulated. In brief, improvement of their own program effectiveness and program refinement are the targeted outcomes of the efforts of the follow-up studies.

Methodology

The methodologies utilized vary from study to study. The authors discuss such methodological issues as study design, variables of the studies, sources of data, kinds of instruments, samples of subjects, data points, data management and data analyses.

Building ownership in the initial study design stage seems to promise pay off in the end. Aubrecht at Ohio State, points out that faculty involvement is essential to insure that "persuasive" data is obtained. One way to do this is through faculty review and agreement. For example, in one study a literature search was first conducted by a nuclear staff for items representing desirable teacher attitudes and behaviors. This was followed by staff brainstorming and organizing the items into "rational" groups before presentation to the entire faculty for their additional items and consensus. Similarly, instrumentation for collecting data relative to the research questions is developed by a program evaluation staff, or by a faculty committee, and then reviewed by the faculty as a whole. In the same vein, in a case where existing instrumentation is employed, a core group assembles the possibilities, which are then reviewed either by a larger group representing the faculty or by the entire faculty. Instrumentation--commercial, or originating on institutional campuses--are named and described in the studies.

The research questions posed by the studies utilize a wide array of variables. Attention most often is focused on the teaching behaviors or competencies exhibited by graduates in their teaching performance and on the contribution of program elements, such as training modules or other components, to meeting graduates' needs and/or developing graduates' competencies. In addition, the importance which raters assign to various competencies is included. Demographic variables are collected. Collecting information on the attitudes of graduates about teaching and being a teacher, on interpersonal relationships and on a variety of personality factors, is common. Characteristics of the school setting as the context for teaching, including the judged difficulty of the school setting, are sometimes a part of the data which are gathered. In one study pupil outcomes and evidence of pupil achievement are requested.

Who does this information come from? It is collected from program graduates while they are students in the program and during their inservice as teachers. Trained observers, who are sometimes college faculty, provide data. Other training program faculty serve as sources of information about the graduates. The college supervisor is cited quite frequently as a data source, as are school building principals and other district administrators and supervisors, and classroom cooperating teachers. Information of an evaluative nature is collected from the pupils whom graduates teach and from the graduates' peer teachers. The representatives of professional organizations are sometimes approached for data. Interestingly, no mention is made of going to the parents of the children in the graduates' classrooms to gain data.

What kind of instruments are used in the follow-up studies to obtain information for program evaluation? Surveys are frequently employed for identifying teacher behaviors and program elements; classroom observations provide descriptions of teaching performance; questionnaires are used to obtain demographic information. Surveys and questionnaires, in addition to inventories (of personality traits), rating scales, and checklists are frequently distributed by mail, though any or all of these are sometimes administered on-site. Classroom observations are done on-site, of course. Other on-site activities are interviews with the graduates, their principals, peer teachers, and cooperating teachers. Finally, evaluation forms are administered to pupils or students of the program graduates. Quite often instrumentation is pilot-tested and revised.

In addition to the above kinds of instruments, the permanent records of graduates reflecting grade point average and other academic information, as well as the National Teacher Examination scores, and other standardized measures contribute to the data base.

Frequently, the sample of the subjects (graduates) is stratified in order to provide an array of variables, such as age, sex, teaching subject area, grade level assignment, and years from graduation. The geographical boundary variable is usually determined by convenience: a fifty-mile radius seems common, though 100 miles is used in one case.

Depending on budget (often the determining factor), data points most often identified are: during the undergraduate student teaching experience, following student teaching, end of first, third, and fifth years of teaching. One of the studies collects data after each of the first five years of teaching, and one study has long-range plans for a nine-year data collection scope.

As expected, the handling of data management and its storage depend on budget. Adams suggests that an important consideration in terms of time (and out-of-pocket cost) is the scoring of instruments by machine or computer. Several studies describe a management and storage system where data files are placed on disks; others use tape for storing data; while computer card decks handle data in such a way as to maintain a computer card for each subject for each year of the study, with data from each instrument.

Data analyses range within studies and across studies from simple descriptive summaries to complex multivariate analyses. In the less sophisticated, information is mainly descriptive; data is summarized in tables and charts with the provision of summary reports. Computer analyses may involve factor analysis on certain data sets, application of t-test, and correlational analyses to determine relationships among variables and analysis of variance to determine differences across years. Pre- and post-analyses of preservice and graduate students are done when data are available. Concerns for reliability and validity are addressed in several of the studies. Reliability of survey and observation instruments and of interview procedures is established and checked; interobserver reliability is established each year. Validity is addressed through pilot testing, through faculty consensus, and through correspondence of responses of graduates, their supervisors, and their peers. Observation data has been used to check the validity of graduates' self assessment.

Costs

Little information is provided about the direct expenses required for conducting follow-up studies. Yet, those who were able to share were very explicit: at Tennessee Technological University the development and implementation of the longitudinal study of graduates was $15,000 per year; the Oregon College of Education 1976 Elementary Teacher Study cost $10,000 for the on-site approach and $2,000 for the telephone-mail design. In the same year, the Oregon Secondary Teacher Study on-site design required only $3,500 in out-of-pocket costs, because college supervisors assumed the on-site visitations. The 1977 Secondary Teacher Study at Oregon College of Education cost \$5,199 in out-of-pocket expense, \$12,461 in contributed time by college faculty and administrators, and \$3,296 in contributed time by the schools. Program evaluation and follow-up studies clearly cost real dollars.

Feedback

Study findings are translated and prepared as feedback to various audiences. The faculties of teacher education programs normally receive a summary report of findings for each year. In addition, at one institution, the faculty receives an oral presentation by the evaluation staff who explain the information and give program recommendations based on the evaluation information. In another institution, feedback is given in a technical report; it is not translated into programmatic implications, and there is little faculty interest, despite the fact that they help collect the data. In contrast, feedback is used by another teacher education faculty to consider defficiencies as they make curriculum decisions. It is used by program managers on another faculty to identify priority areas needing improvement. Invitations for special analyses and interpretations are commonly made to faculty members and students. In some cases, faculty share and interpret feedback to students.

At some institutions students receive their own summary data and the total group data each year. In other institutions, group data are shared only at the end of the five years. If preservice data are gathered, they are sometimes used with students for individual counseling.

Who else is given feedback? The administrators of teacher education programs are an audience. Feedback is used with administrators as a means of seeking program support from the college of education. Another feedback target are prospective students who are being recruited for undergraduate programs. Public school educators who hire graduates are another audience, as are local, state, and national accreditation groups. Principals who participated in the study are given reports. Curriculum re-design groups and the dean's staff also receive summaries.

Problems

In large measure, the problems cited in the studies can be grouped into two areas. The first area is computer management of data: coding data, key punching, programming, computer services, computer operations. Related to this problem is the conduct of data collection and the drop-out rate of the sample from year to year. The second area is the lack of resources which are needed in order to look at relationships between variables and overall competency/effectiveness of graduates, to collect pupil achievement scores, to measure classroom climate, to do intensive case studies. Resources are needed for more follow-up work (travel, mail, phone), for doctoral student support to do research, for the enormous task of collecting and analyzing on-site data, and for tailoring reports to fit various audiences.

There is disappointment in the lack of full use of data by faculties and the slow rate of subsequent program change and improvement. There is concern for a more adequate system for evaluating persons with advanced degrees. There is concern as well for the validity of the evaluation model and the appropriateness of variables, for the reliance on high inference judgments, and for the questionable validity of mail/telephone data.

Despite the concerns, the questions, and the less-than-hoped for conditions, follow-up studies will continue to be done by the investigators in these institutions.

Implications for Future Work

Considering what has been learned about teacher training programs, program graduates, and how to conduct follow-up program evaluation, what issues now need to be addressed? The findings hold implications for (1) future follow-up studies, (2) research in teaching, and (3) research in teacher education.

Highlights of the studies' results have been provided by authors. These findings are presented through tables, charts, graphs, explanatory comments in descriptive and statistical formats. Study findings in conjunction with the problems discussed by the investigators provide some sugggestions of what needs to be considered next.

Implication #1: Program evaluation and follow-up studies have real costs.

Program evaluators cite the need for resources, for data collection and analysis, and for tailoring reports to different audiences. The costs of these studies have a high dollar price. Costs are computed not only in terms of dollars but also in the time it takes to train institution faculty and the public school staff.

Implication #2: The findings are not being used.

Most authors pointed out that the results of their studies are not being used sufficiently; the impact of the feedback is disappointing; the purpose of the study is not being served. It has been suggested elsewhere that changing a college is like reorganizing a cemetery (attributed to S. Freud by Hall & Jones, Competency Based Education, 1976). Still, new ways must be found or created for initiating and designing reform. How can this happen? Will getting faculty more involved, more committed be effective? A suggestion was made by one author to get faculty and evaluators together to interpret data and decide on programmatic changes. How can this strategy be used? Are there ways to energize decision-makers? The presentation and use of results are clearly in need of fresh thinking and new ideas.

Implication #3: Institutional administrators must actively support the studies.

A key to the solution of the first and second issues is the need for developing administrative procedures to encourage and/or support the changes which are planned or designed. Results are shared with college of education administrators to gain support for programs. After the sharing, what happens next? What effective administrative actions can be utilized to support the implementation of changes once they have been designed?

Implication #4: Graduates don't seem to fit the program goals.

It has been suggested that program revisions should be based on the needs and abilities of graduates; the results show that graduates seem to need some help. This idea is confounded by the suggestion that changes in admission procedures may be required -- to obtain a different breed of student. Evaluating programs means getting an intimate look at the students; much is learned about the individuals while learning about the programs. Is this a double-edged sword? Do you tailor the program to the student or change the student candidate selection to fit the program? A knotty problem. Might modifications in each offer maximum results?

Implication #5: It is not clear exactly when and how a teacher establishes his or her style.

One study reports that over four years of teaching, graduates have changed very little. Another finding indicates that over time, changes are revealed in teaching performance. More work needs to be done to answer the question: At what critical point(s) is a teacher's style developed? Then comes the question of how can a teacher's style be modified? College or public school-based inservice may be desirable at this juncture. If so, what kind? How delivered?

Implication #6: How should teacher education attend to different school contexts?

The research on teaching clearly indicates the importance of context in effective teaching. That is, teacher behaviors which are effective in one setting (grade level, subject matter, type of student, etc.) may not be effective in another setting. Should teachers be trained for a specific context? If generic teaching strategies are not generic, but are related to particular context variables, how can teacher training be managed to accommodate this phenomenon? Additionally, what effect does classroom or school "context" have on teachers? As Schalock inquires, are there more "difficult" schools? In his report, graduates in "difficult" classrooms were judged less competent in teaching performance.

Implication #7: What are the essential elements of school/college collaboration?

Collecting data on-site in the schools has been described as resulting in strong benefits for "public relations" between the institutions. These visits may well provide the basis for developing dialogue and collaboration with field sites. There are a host of reports extant in the literature which address "collaboration" between teacher training institutions and the public schools as laboratories for education students. Yet, these reports by and large are unclear as to what factors are important for developing collaboration or how to go about it once the significant variables are recognized. Much can be gleaned from follow-up studies and the experiences of follow-up evaluators to inform such future collaboration. Perhaps follow-up studies can provide the "seed" for developing closer, more problem-oriented relationships.

Implication #8: Study results could be used to initiate staff development.

Evaluation of training programs and its results may serve to diagnose the needs of the college faculty for staff development. The findings may also provide relevant information for the needs and improvement of public school faculty. What are the effective means of encouraging and supporting staff renewal in universities? in schools?

The fabric of teacher education is not whole cloth; it is woefully incomplete. What do we know about the content or what should be included in preservice and continuing inservice teacher education programs? How should programs be delivered? Which delivery processes are most effective in which contexts? What do we know about how to support professionals in collaborative modes so that the most relevant programs and delivery systems are provided in congruence with the special characteristics of adult learners? These evaluation studies are concerned about these questions. They have begun to address them.

Because the authors have been involved in conducting follow-up studies for three to eight years, they have gained experience and much expertise. Their papers provide important "how-to-do-its" for designing and managing effective studies. The authors respond to the frequent requests they receive for information on follow-up studies. The reader is invited to solicit further information from the contributors. A list with addresses is provided at the end of this volume for that purpose.

LIST OF CONTRIBUTING AUTHORS

Ronald D. Adams
Director, Office of Educational Research
College of Education
Western Kentucky University
Bowling Green, Kentucky 42101

Harley K. Adamson
Weber State College
3750 Harrison Blvd.
Ogden, Utah 84408

Richard I. Arends
Division of Teacher Education
University of Oregon
Eugene, Oregon 97405

Judith D. Aubrecht
251 North Franklin St.
Delaware, Ohio 43015

Jerry B. Ayers
Associate Dean
College of Education
Tennessee Technological University
Cookeville, Tennessee 38501

Patricia E. Blosser
Faculty of Science and Mathematics Education
College of Education
283 Arp Hall
1945 North High Street
The Ohio State University
Columbus, Ohio 43210

Caseel D. Burke
Weber State College
3750 Harrison Blvd.
Odgen, Utah 84408

James M. Cooper
College of Education
University of Houston
Houston, Texas 77004

J. H. Garrison
Oregon College of Education
Monmouth, Oregon 97361

G. R. Girod
Oregon College of Education
Monmouth, Oregon 97361

Robert W. Howe
College of Education
The Ohio State University
Columbus, Ohio 43210

Howard L. Jones
College of Education
University of Houston
Houston, Texas 77004

K. H. Myers
Oregon College of Education
Monmouth, Oregon 97361

Mikell O'Donnell
110 Arp Hall
1945 North High Street
The Ohio State University
Columbus, Ohio 43210

Robert Randall
College of Education
University of Houston
Houston, Texas 77004

Kevin Ryan
Academic Faculty of Educational Foundations and Research
121-C Ramseyer Hall
20 West Woodruff Avenue
The Ohio State University
Columbus, Ohio 43210

H. D. Schalock
Research Professor
The Teaching Research Division
Oregon State System of Higher Education
Monmouth, Oregon 97361

Wilford A. Weber
College of Education
University of Houston
Houston, Texas 77004